

AD\_\_\_\_\_

Award Number: DAMD-17-03-2-0001

TITLE: Advanced Technologies in Safe and Efficient Operating Rooms

PRINCIPAL INVESTIGATOR: Adrian E. Park, M.D.

CONTRACTING ORGANIZATION: University of Maryland Medical Center  
Baltimore, MD 21201

REPORT DATE: February 2006

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command  
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;  
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. <b>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</b>					
1. REPORT DATE 01-02-2006		2. REPORT TYPE Annual		3. DATES COVERED 1 Mar 2005 – 28 Feb 2006	
4. TITLE AND SUBTITLE  Advanced Technologies in Safe and Efficient Operating Rooms				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER DAMD17-03-2-0001	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)  Adrian E. Park, M.D.  Email: <a href="mailto:apark@smail.umaryland.edu">apark@smail.umaryland.edu</a> ; <a href="mailto:gmoses@smail.umaryland.edu">gmoses@smail.umaryland.edu</a>				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)  University of Maryland Medical Center Baltimore, MD 21201				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited					
13. SUPPLEMENTARY NOTES Original contains colored plates: ALL DTIC reproductions will be in black and white.					
14. ABSTRACT Work on this award continues to focus on three primary areas: Informatics, Simulation and Smart Image. In the area of Informatics, a Meta analysis of commercially available positioning technologies has been completed. A sampling of these technologies was installed in a lab environment and extensive testing was completed. The Simulation effort continues to build The Virtual Human, having developed a working model of normal esophageal function. The Smart Image research has demonstrated the ability to fuse multi-source images, creating a real-time 3D view for the laparoscopic surgeon. Work continues on refining these images.					
15. SUBJECT TERMS No subject terms provided					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	273	19b. TELEPHONE NUMBER (include area code)

## Table of Contents

<b>Introduction.....</b>	<b>4</b>
<b>Body.....</b>	<b>5</b>
<b>Conclusion.....</b>	<b>11</b>
<b>References.....</b>	<b>12</b>
<b>Appendices.....</b>	<b>15</b>

## **Introduction**

This Annual Report covers the research period ending January 1, 2006 for the Cooperative Agreement Award No. DAMD 17-03-2-001. Under the direction of Dr. Adrian Park, MD., Principal Investigator, research has continued in the three primary focus areas which were defined in the approved contract modifications. These are OR Informatics, Simulation and Smart Image.

### **Research Objective 1: OR Informatics: Process Simulation**

The focus of this research effort has been the identification and evaluation of commercially available positioning technologies for potential use in the perioperative environment.. It is hypothesized that data gathered by these devices could be used to populate a workflow engine that would initially help define best practices and ultimately assist in managing the perioperative workflow. It is hypothesized that by insuring the necessary resources converge in the right place at the right time, patient safety would be improved, efficiency would be enhanced and the level of chaos in the perioperative environment would be reduced.

### **Research Objective 2: Simulation**

This research effort continues the development of the Virtual Human that will enhance medical education, promote patient safety and allow for training of clinicians without the use of a “live” patients. The focus has been on the continue encoding of ontological concepts, properties and concepts in a software framework that allows user interactions with the Virtual Human.

### **Research Objective 3: Smart Image**

This effort continues to focus on the limited visualization problems in minimally invasive surgeries by merging various images and technologies to allow the display of a converged or Smart Image. This technology has been demonstrated in laparoscopic surgery on animals and work continues to enhance the effectiveness of such images.



## **Annual Report**

The operating room and perioperative process is the core of intense care provided for many patients. The operating room forms the nucleus of mobile military hospitals. Whether found in civilian or military healthcare, the operating room is a high-cost and high-risk care environment. Many types of cases, multiple surgical specialties, growing use of technology, and the sheer number of elements makes the operating room the most difficult environment to manage in healthcare.

The Operating Room of the Future Project at the University of Maryland Medical Center continues to provide analysis and data on the novel use of communications and other technologies achieving and maintaining situational awareness in high velocity operating rooms. While the rapid evolution of technologies requires constant re-evaluation of the technologies being evaluated and their potential for deployment, the focus continues to be on work that coincides with the DoD Telemedicine Science and Technology Plan in the operational capability areas of 1) Joint Medical Readiness; 2) Battlespace Medical Awareness; and 3) Effective Employment of Medical Forces. The common operational challenges for both civilian and military healthcare in the future addressed by this work include overcoming geographical distances, continual training requirements for medical forces, and high volume, chaotic communications environments in the field and fixed site medical facilities.

The University of Maryland Medical System in conjunction with the University of Maryland School of Medicine and others, continues to work on three primary areas. Progress in these areas is detailed in the following report.

In addition to the direct research, the ORF Project sponsors an Annual Retreat, designed to bring together researchers and others involved in this arena to exchange ideas, to identify synergies and to identify partnering opportunities. The 3<sup>rd</sup> Annual Retreat will be held in June and will be the largest yet, with nearly 100 invitees.

## **Informatics**

The focus of this initiative has been on the identification and assessment of commercially available locational technologies, their suitability for deployment in the perioperative environment and how they could be used to populate a work flow engine. This work has been documented in “A Survey of Location Technologies and their Application in Health Care”, attached as Exhibit 1-A. As part of this assessment, extensive testing was conducted to validate the actual performance of the various technologies. The test plan and findings have been documented in “UMMC Indoor Positioning System Evaluation Facility”, attached as Exhibit 1-B. A draft of an article detailing the actual test results is attached as Exhibit 1-C.

As identified in the attachments, the development, marketing and deployment of RFID and IPS for healthcare applications is in its early stages. The market is divided among a number of competing technologies based on non-standard and incompatible technologies. There have been many pilots, though the reporting on these projects provides little quantitative information on the performance or earned value of the systems.

However, these technologies continue to evolve and we anticipate that one or more of them will be effective in populating a workflow engine. Initially, this data will tell us what is happening; the ultimate goal is to define best practices and to issue appropriate alerts when events are not on track. By insuring the necessary resources converge in the right place at the right time, patient safety will be improved, efficiency will be enhanced and the level of chaos in the perioperative environment would be reduced.

In addition to the direct research, the Informatics team has established a relationship with CIMIT to explore and exploit how our efforts can benefit both organizations.

# **Smart Image: Development of Continuous CT-Guided Minimally Invasive Surgery**

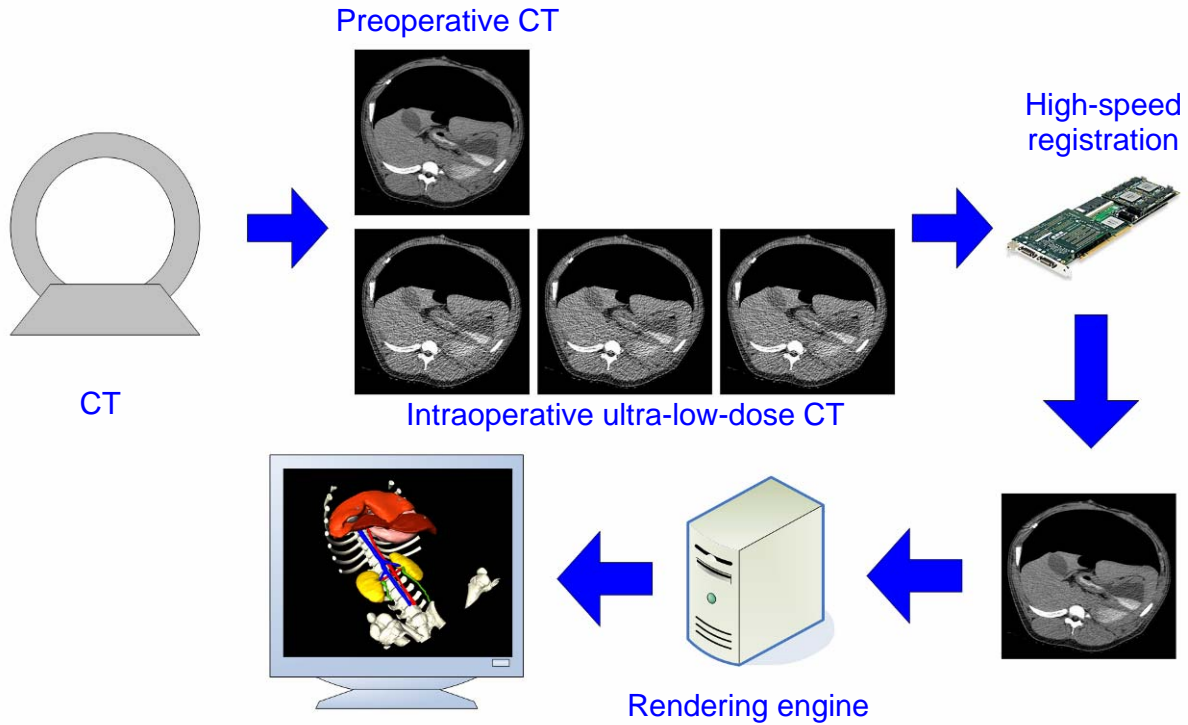
## **Introduction**

Minimally invasive laparoscopic surgeries present an attractive alternative to conventional open surgeries. In these procedures, internal anatomy is accessed through a few small ports in the patient's skin rather than through large incisions. Compared with conventional open surgeries, minimally invasive surgeries have been shown to lead to improved outcomes, less scarring, and significantly faster patient recovery.

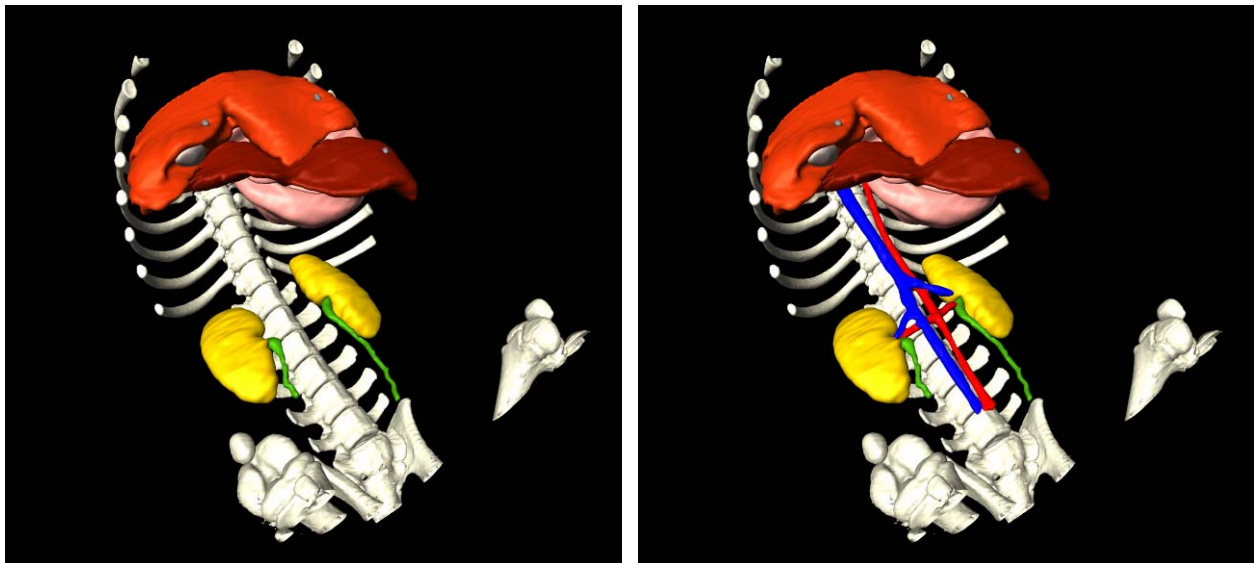
Despite the success of minimally invasive surgeries, the laparoscopes used to guide these procedures continue to be a limiting factor. They provide a relatively flat representation of the 3-dimensional (3D) anatomy and can display only the most superficial surfaces. Moreover, visualization of the vasculature remains a challenge. Laparoscopes are by their nature limited in that they cannot see inside a structure before dissection—a long-standing need of laparoscopic surgeons. Our research is addressing this unmet need by using continuous multislice computed tomography (CT) for both intraoperative visualization and surgical navigation.

## **Approach**

CT can provide enhanced intraoperative visualization far superior to that of laparoscopes, but radiation exposure to the patient and the surgeon remains a concern with the use of continuous CT. A major thrust of our research, therefore, is to design, develop, and test the following dose reduction strategy and incorporate it into our proposed continuous CT-guided surgical navigation system.

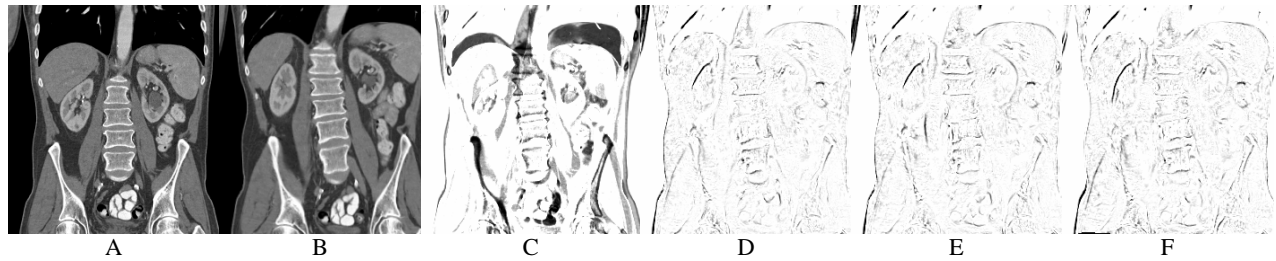


**Figure 1:** Schematic explaining the concept behind CT dose reduction through image registration.



**Figure 2:** Vessels cannot be visualized if intraoperative CT is rendered (L) but can when warped preoperative CT (R) is rendered.

Our dose reduction strategy, which we have been developing for the past 2 years, is illustrated in Figure 1. The concept behind this strategy is to acquire a standard, high-quality CT (collected at 200–250 mAs) image preoperatively (after creation of the pneumoperitoneum) and scan the dynamic operative field using ultra-low-dose CT (at 10 mAs or lower) once surgery begins. Using high-speed, nonrigid 3D image registration (warping) techniques developed by our group



**Figure 3:** Registration errors with low-dose CT: (A) original image; (B) image after known realistic deformation; (C) corresponding difference image; (D–F) difference images after registration with low-dose CT (at 200, 50, and 10 mAs, respectively) with respect to the deformed image.

[1–5], we will rapidly register the preoperative CT image to the ultra-low-dose intraoperative CT image. Moreover, we plan to repeat registration for each new 3D CT image arriving at 8 frames/second. Warped preoperative CT images, which show the intraoperative anatomy accurately, will then be substituted for the ultra-low-dose, poor quality CT. The resulting images will be 3D rendered and presented to the surgeon from the perspective of the laparoscope. The computer-generated imagery also will be superimposed on the laparoscopic video to create augmented reality views.

A significant added benefit of this approach is intraoperative visualization of the vascular anatomy. Vessels can be contrast enhanced in the preoperative CT. Rendering the warped preoperative CT to visualize intraoperative anatomy then retains the enhanced vessels (see Fig. 2). Note that vessels are absent if the intraoperative CT is visualized directly. Moreover, because contrast agents cannot be injected continuously during surgery, our approach provides a unique solution to visualize vessels intraoperatively.

## Significant Achievements

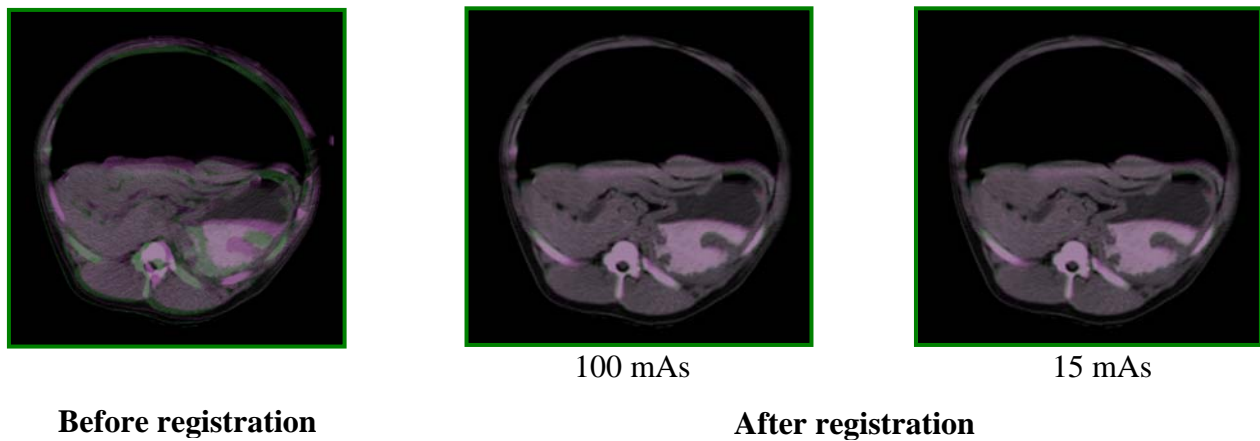
(1) Determination of low-dose threshold.– We have investigated dose reduction through the aforementioned strategy by registering a standard CT image with a low-dose CT image (dose level, 200–10 mAs). Note that the dose cannot be made arbitrarily small, because this could cause registration to fail or could affect registration accuracy, directly and negatively affecting intraoperative targeting accuracy.

The original image (standard CT) and a target image with known deformation are shown in Figures 3A and 3B, respectively. Figure 3C shows the starting misalignment between these 2 images, Figures 3D–F indicate the difference after nonrigid registration at doses of 200, 50, and 10 mAs, respectively. Visually correct registration of the standard CT image (preoperative) with the target image (intraoperative) at various low doses (evident from the reduced features in the difference image) demonstrates the feasibility of nonrigid registration at low CT doses.

A quantitative method to evaluate the accuracy of nonrigid registration with low-dose CT has been described previously [6–7]. The underlining principle of this method is to start with a low-dose image with known deformation and then compare the registration outcome with the known

reference solution. The nonrigid registration achieved an average registration accuracy of 2.25 mm at the lowest dose (10mAs), with isotropic CT image resolution of 1.5 mm.

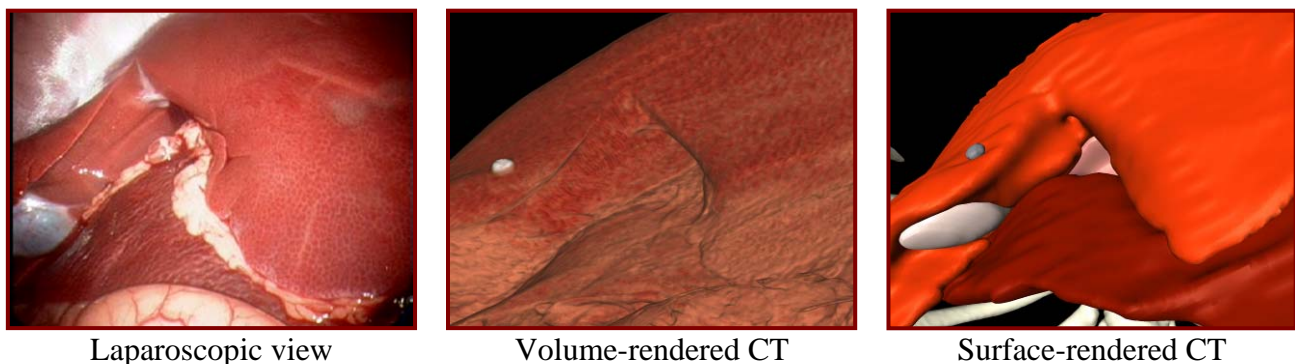
Figure 4 shows the result of nonrigid registration between pre- and intraoperative images in an animal model. Qualitative evaluation indicates reasonable registration accuracy even at 15 mAs. We are currently investigating quantitative methods to evaluate intraoperative registration accuracy.



**Figure 4:** Nonrigid registration in an animal model: fusion of pre- and intraoperative CT.

(2) Tool tracking. – We have tested the feasibility of tracking the laparoscope and correlating its location/orientation in the CT coordinates and proven the concept. A dedicated tool tracking system for the CT room, which also permits temporal synchronization among various devices, will be created for future experiments.

(3) CT-generated views. – We have developed methods to produce static CT-generated views corresponding to laparoscopic views. Figure 5 demonstrates this capability for a laparoscopic frame. These methods have been also extended to provide dynamic CT-generated views.



**Figure 5:** Intraoperative visualization

#### (4) Publications and presentations

- Dandekar O, Walimbe V, Siddiqui K, Shekhar R. Image registration accuracy with low-dose CT: How low can we go? Presented at IEEE International Symposium on Biomedical Imaging, April 2006; Arlington, VA. (Reprint in Appendix)
- Shekhar R, Dandekar O, Weiner M, Malloy P, Mezrich R, Park A. Development of continuous CT-guided minimally-invasive surgery. Presented at Biomedical Imaging Research Opportunities Workshop, February 2006; Bethesda, MD. (Poster in Appendix)

### **Conclusions**

Our deliverable remains a prototype of a continuous CT-guided surgical navigation system. Our goal is show its real-time operation once CT images have been reconstructed. An online system ready for clinical trials will be the focus of a follow-up project. This will require integration of our prototype with a CT scanner and a joint partnership with a CT vendor towards creation of a surgical CT scanner.

See Article “Image Registration Accuracy with Low-dose CT; How Low Can We Go?”,  
attached as Exhibit 2-1.

## References

1. Walimbe V, Shekhar R. Automatic elastic image registration by interpolation of 3D rotations and translations from discrete rigid-body transformations. *Med Image Anal.* Accepted for publication in 2006.
2. Castro-Pareja CR, Daly B, Shekhar R. Elastic registration using 3D chainmail: application to virtual colonoscopy. *Proc SPIE Med Imaging.* 2006; in press.
3. Shekhar R, Zagrodsky V, Castro-Pareja CR, Walimbe V, Jagadeesh JM. High-speed registration of three- and four-dimensional medical images by using voxel similarity. *RadioGraphics.* 2003;23:1673–1681.
4. Castro-Pareja CR, Jagadeesh JM, Shekhar R. FAIR: a hardware architecture for real-time 3-D image registration. *IEEE Trans Inform Technol Biomed.* 2003;7:426–434.
5. Castro-Pareja CR, Shekhar R. Hardware acceleration of mutual information-based 3D image registration. *J Imaging Sci Technol.* 2005;49:105–113.
6. Dandekar O, Walimbe V, Siddiqui K, Shekhar R. Image registration accuracy with low-dose CT: how low can we go? Presented at IEEE International Symposium on Biomedical Imaging, April 2006; Arlington, VA.
7. Shekhar R, Dandekar O, Weiner M, Malloy P, Mezrich R, Park A. Development of continuous CT-guided minimally-invasive surgery. Presented at Biomedical Imaging Research Opportunities Workshop, February 2006; Bethesda, MD.



## **MVP: The Maryland Virtual Patient Progress Report**

Bruce E Jarrell, MD, Department of Surgery, University of Maryland School of Medicine  
Sergei Nirenburg, PhD, Marjorie McShane, PhD, Stephen Beale, PhD, Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County

In the past year, our team has made significant progress in the areas of design, knowledge elicitation, knowledge representation, system development and environment development for MVP, the Maryland Virtual Patient.

### **Creation of the Conceptual Infrastructure**

- We invented the notion of the “Double Agent”, which combines physiological and cognitive capacity in the virtual patient and other virtual agents in the system. For an overview, see **Appendix 3-A** (“Double Agent: An Environment for Automatic Tutoring of Medical Students Using Simulation Involving a Combination of a Physiological Software Agent, a Cognitive Software Agent, Software Agents Representing Members of the Medical Team and a Combination of Human and Software Agents for Mentoring”) and **Appendix 3-B**, the attached Power Point presentation entitled **The Double Agent**.
- We created **knowledge elicitation methodologies** and a **conceptual framework** for modeling diseases, diagnostics and treatments. The methodology and framework grew out of our knowledge elicitation sessions, meaning that the knowledge itself, in conjunction with the needs of the project, determined, in large part, the form in which we recorded the knowledge.
- We researched other extant approaches to medical modeling, simulation and virtual patients and formally distinguished our approach to virtual patients – which we call “Maryland Virtual Patients” or MVPs – from all other extant approaches. The distinctions are described in two papers (see **Appendices 3-B** and **3-C**).

### **Modeling of Basic Physiology**

- We expanded and refined the scripts for normal esophageal function, including swallowing, that were created in year 1 such that they supported the modeling of two diseases (see below).
- We expanded the ontology and accompanying lexicon to include the concepts needed for modeling the two new diseases.

### **Modeling of Diseases, Diagnostics and Treatments**

- We modeled two diseases, achalasia and GERD, and implemented instances of MVPs (“Maryland Virtual Patients”) with these diseases. For GERD, we implemented 6 of the 7 possible disease paths (proximal GERD is not yet handled). See **Appendices 3-D** and **3-E** for descriptions of these diseases (these descriptions are project-internal documentation). These appendices include the questionnaires that MVP authors need to fill out to create new patient instances as well as sample MVP instances.

## Simulation

- We enhanced the simulation engine from Year 1, improving the handling of basic physiological scripts and incorporating from scratch the handling of diseases, diagnostics and treatments.
- We worked on incorporating a visualization component using AnyLogic and made progress in graphically rendering swallowing and the effects of swallowing of certain disease processes, like tumor growth. However, we are not continuing to pursue that aspect of the work due to the unexpectedly large amount of resources it required.

## Demonstration Version of the MVP System

- We created a demonstration version of the MVP system that provides a user interface and four MVP instances – 2 GERD patients and 2 Achalasia patients. A user can select, diagnose and treat these patients. Apart from providing insight into the ultimate student experience when using this system, the interface also provides a view of the actual physiological properties of the patient during the entire process, with those properties that are important to the given disease highlighted. This view will not be available to students in the in the tutoring scenario but is available to physician-teachers and developers.

## Tutoring

- We have begun developing a conceptual infrastructure and knowledge elicitation methodology for tutoring. The basic approach is to teach the system to understand what an expert's next move would be during a simulation by attaching certain types of “preconditions for good practice” to diagnostic procedures, treatments, etc. We are developing an inventory of properties and values that will inform the simulator's evaluation function whether or not a given student move is correct or whether, by contrast, the student has veered off the path of good practice. This all amounts to the creation of so-called **workflow scripts**.
- We have done extensive background reading on other simulation and tutoring systems, with CIRCSIM-Tutor being of particular note. We will use their experience in tutoring as a springboard for our tutoring module.

## Dissemination of Results

- As mentioned above, we have written two papers, one of which will be presented as a poster at the FLAIRS conference in May 2006 (**Appendix A**), and the other of which will be a book chapter (**Appendix B**).

## **Appendices**

## **Appendix 1-A**

# **A Survey of Location Technologies and their Application in Health Care**

**29 April 2005**

**Bill Bame, UMMS  
Duncan Clarke, Fremont Associates  
Mary Rose Connolly, UMMC  
Mike Minear, UMMS  
Adrian Park, UMMC**

## Table of Contents

1. Introduction.....	3
2. Technology Description.....	3
a. Definitions .....	3
b. RFID Technologies.....	5
c. Issues in Healthcare .....	7
d. Technologies .....	8
e. Standards Activity.....	9
f. Proposed Selection Criteria .....	10
3. Applications.....	12
a. Patient Identification.....	12
b. Automated Verification .....	13
c. Infant Security.....	14
d. Workflow and Equipment Monitoring and Management.....	14
4. Vendor Review .....	15
e. Low-Level RFID Technology .....	15
f. Healthcare Oriented RFID and IPS .....	15
g. IPS Enabled Workflow Solutions .....	17
h. Healthcare Oriented RFID Materiel .....	17
5. Literature Review .....	18
a. Refereed Journals.....	18
b. Web Sources .....	19
c. Engineering/Computer Science Literature.....	19
6. Summary .....	20

## 1. Introduction

Passive technologies for enunciating the identity and location of equipment, supplies and people are currently an important topic in the healthcare industry. Of particular interest are so-called Radio Frequency Identification (RFID) technologies, already heavily used in high-volume retailing, that involve affixing small radio transceivers to key assets to track resource flow. Though much of the discussion is driven by the technology, the underlying problems being addressed are not new: How can we track the location of equipment, supplies and people to address needs like efficient utilization, loss prevention and accuracy of dispensing, treatment and billing?

In this document we attempt to step back from the current technology-centric focus of the trade press and assess the application of location technologies in the health care field. We begin in Section 2 by assessing location and identification technologies broadly, but with a particular emphasis on RFID and other indoor positioning technologies, since they are relatively new to the health care setting. Section 3 provides an overview of recent implementation projects in health care from the trade press and journals. Section 4 provides a sampling of key vendors in the field with brief technology descriptions and references to past, ongoing and planned projects incorporating their technologies. Section 5 reviews relevant literature. And finally, Section 6 will close this document with summary remarks.

Our intended audience is decision makers in the health care industry who want a deeper understanding of the underlying technologies with an emphasis on applications. The technology presentation will focus on the relative strengths and weaknesses of various technologies, and the survey of applications will emphasize novel early adopter implementation projects.

For readers who worry that they are hopelessly behind in implementing RFID and allied technologies we offer a bit of reassurance. Industry estimates place the number of hospitals using RFID for tracking at less than 1% of the domestic total, and most of those are prototype projects at academic medical centers. There is still plenty of time to be an early adopter of this technology.

## 2. Technology Description

This section presents an overview of location technologies with an emphasis on RFID and other indoor positioning technologies. We begin by defining concepts and terms that will be important throughout the presentation. Then we present a summary of RFID and indoor positioning technologies. Next we assess existing technologies that could be augmented by or replaced by RFID. Finally, we close with an assessment of issues of special interest when applying RFID or indoor positioning technologies in the health care environment.

### a. Definitions

*Perioperative Environment*—We shall adopt as our scope the perioperative environment, which we will interpret broadly to include all functional areas from initial scheduling through discharge (of ambulatory surgery patients) or placement in a hospital bed following first-stage postoperative recovery. Functional areas covered by this definition include OR scheduling, pre-admission screening, preparation the morning of surgery, the OR's, central sterile processing, first-tier postoperative recovery, and either second-tier postoperative recovery and discharge (for ambulatory surgery patients) or placement in a hospital bed.

*The location problem*—Within the perioperative environment the location problem includes determining the location of a person, a unit of equipment, or a supply item automatically and in real-time. This location information can be used in real-time to facilitate work activities, archived to enable historic analysis of movements of resources, and to create decision models to support process improvement efforts..

Note that different applications require the location problem to be addressed with different levels of precision. In bulk retail shipping applications (the goal of Wal-Mart, for example) it is often sufficient to record the point in time a particular resource, such as a pallet or a case of retail products, passes through various portals or locations (warehouses, retail stores, trucks, et cetera). Many applications that are envisioned for the perioperative environment would require continuous monitoring of resource position anywhere within a large medical center campus, with position reported to the nearest square foot.

*Radio Frequency Identification (RFID)*—There are a number of distinct technologies referred to casually as “RFID,” despite the fact that some don’t use radio frequencies at all. To avoid getting bogged down in issues of terminology, we will accept this broad definition of RFID that encompasses systems consisting of mobile transmitters and fixed or semi-fixed receivers used wirelessly and without human intervention to identify the transmitter to the receiver. Note that this definition specifically excludes applications based on technology utilizing card swiping, bar codes, video surveillance, periodic physical connection to a base station, and human transcription, among others.

*Indoor Positioning System (IPS)*—IPS is a term often used interchangeably with RFID, although they are not synonymous. The term “IPS” emphasizes the problem being solved rather than the technology used. In broad terms an indoor positioning system is responsible for determining the location of resources in an indoor environment. The name was chosen for its contrast with the Global Positioning System (GPS), which solves an analogous problem for resources in the outdoor environment.

*Wired Networks (802.3)*—Many solutions to the location problem rely on existing network infrastructure to carry data from a collection of remote readers to a central processing node. This infrastructure is variously referred to as a (1) wired network, (2) 802.3 network, (3) Ethernet, or (4) LAN. These networks are based on physical connections between sender/receiver pairs.

*Wi-Fi Networks (802.11)*—Some solutions to the location problem make use of standard wireless network technology for their basic approach to location, as a connectionless entry-point to the existing wired network, or both.

*System Architecture*—Solutions to the location problem will typically take the form of systems that integrate transmitters, receivers, networks, a central data repository, application software and user workstations. A *closed system architecture* will be a single vendor-supplied system consisting of all of the components needed to create an application-specific end-to-end solution. An *open system architecture* is a loosely integrated collection of components that can supply location data in a generic form for integration with legacy database and reporting systems, or new or custom systems provided by alternate vendors.

An open system architecture is typically designed to expose certain classic interfaces, including (1) the *device interface* for interrogating readers, (2) *middleware* interfaces for filtering, aggregating, collecting, etc., raw device data and forwarding it to other application programs, (3) a *data management* interface for connection to

standard database systems, (4) *application* program interfaces for interacting with application software, and (5) *presentation* interfaces for displaying and manipulating information through an interface such as a web browser page or windows application program.

**Workflow**—Historically, work has been divided into step-by-step functions, with functional areas created to manage the completion of each function. As customer expectations for timeliness, responsiveness and transparency have changed, the emphasis has shifted away from functional management, toward an emphasis on workflow management. In workflow-based management the emphasis is on the smooth and efficient receipt, performance, and forwarding of work items from ordering through to delivery.

**Workflow Engine**—Workflow-based businesses are often supported by a tool called a workflow engine. Workflow engines are software applications that understand the entire end-to-end processing needs to complete an order, and manage the flow of tasks through the system by prompting workers to process work in accordance with established ordering and timeliness constraints. This elevates the status of multi-step jobs so that their completion is no longer the byproduct of sequential functional steps. Instead, now their overall timely and efficient completion becomes the main goal of the overall process.

## b. RFID Technologies



**Figure 1: RFID Hardware**

The simplest, lowest cost RFID implementations found commonly in retail environments consist of tags containing a low-power radio transceiver and a small amount of memory, and readers consisting of a transceiver capable of communicating with the tags and a computer interface. Figure 1 is an image of a low cost Texas Instruments RFID system with three sample tags and their reader.

The tags in this system are *passive*, meaning that they require no batteries or external power source—the power required for operation comes from the RF energy emitted by the reader. A consequence of passive technology is that the tags can operate only in close proximity to a reader; they cannot spontaneously transmit a signal. The maximum usable distance between the tag and the reader is termed the *read range*. Read range will vary depending on the tag size, frequency, and transmitter power of the reader. RF-reflective materials in the vicinity of the tag or reader can also impact read range. Typical read ranges are from 20 to 200 cm, depending on the system configuration and environment.

Another basic feature of RFID is the ability of the tag to store data in memory and transmit the stored value to a reader. The amount of data that can be stored in a tag and the amount that can be transmitted to a reader in a single interaction varies. Low-cost tags like those shown in Figure 1 typically hold about 256 bits. Larger, more expensive COTS tags can hold up to 64,000 bits. As a point of comparison, a simple nine character alpha-numeric patient identifier would require 72 bits of storage, or slightly more than 25% of the capacity of a 256 bit tag. 64,000 bits of memory is sufficient to hold two to three pages of typed text.

Another attribute differentiating tag memories is the contents of memory. A *read-only* tag can be written



**Figure 2: RFID Wristband**

ability to rewrite the exactly once, after



which the identifying information stored inside remains immutable for the life of the tag. *Read-write* tags can be written repeatedly, and are typically certified to operate through at least 100,000 write operations. Read-only tags are used in applications requiring the ability to uniquely identify the tag throughout its life. Read-write tags offer the increased flexibility of being able to augment the original identifying information over time, temporarily maintain off-line state, or completely reprogram the tag for use in a new environment.

Though the memory capacity and read-write capabilities of smaller, lower-cost tags is severely limited, this doesn't necessarily preclude their use in applications requiring the association of voluminous or dynamic data with a unique identifier. A unique 256 bit identifier is ample key information to uniquely identify a tag, allowing additional data and dynamic content to be stored online in low-cost disk storage accessible by tag identifier. Figure 2 illustrates a patient wristband with an embedded RFID tag like those shown in Figure 1. These tags typically encode only a patient identification number, but that number can be used as a key to look up any associated patient records.

When multiple tags will be read in close proximity to one another, another important characteristic of RFID systems is their ability to tolerate *collisions*, *i.e.*, simultaneous replies generated by multiple tags to a single request from a reader. Older RFID technologies required that at most one transceiver tag could be within the operational field of a reader during a read cycle. Current technology incorporates collision tolerance that can accommodate hundreds of tags in the reading zone. The actual degree of collision tolerance of the tags and the effective data transmission rate will depend on a number of factors including the specific technology used and the operating environment.

A dual problem of *reader collision* also exists. That is, when multiple readers are located such that their reading zones overlap, a tag may be interrogated by two readers simultaneously. Reader collision can be a problem if the system design does not account for overlapping read zones. However, reader collision can also be exploited as an enabling technology for determining physical location, as we shall see below.

Note that up to this point we have only talked about technologies that operate by bringing a tag into close proximity with a dedicated reader. Such technologies can address very limited forms of the location problem, such as determining when a tag passed a threshold, or what tags are currently in close proximity to a particular reader. Thus, such RFID systems should not be confused with general purpose Indoor Positioning Systems.

A first step toward the creation of a more general IPS is to extend the range of the tags. This is done by employing *active* tags, rather than passive tags. Active tags incorporate a battery or connect to an external power source to provide sufficient energy for spontaneous long-range read or read-write operations.

Figure 3 depicts an active RFID tag Ekahau, Inc. Note how the addition of a power transceiver has changed the profile of Compare these small pager-sized devices to nominal thickness shown in Figure 1. available, though none come close to the the passive RFID tags.

produced by battery and high the device. the tags with Smaller tags are minimal footprint of



Figure 3: Active RFID Tags

The increased read range of an active tag enables the creation of RFID systems on a much larger scale. A typical application of such tags is the tracking of overseas shipping containers in a container ship or dockyard, or the tracking of railroad cars in a rail yard. As the power and functionality of an active tag increases, the line between RFID and traditional radio or satellite based tracking systems becomes blurred.

Another feature enabled by active tag technology is the continuous transmission of identity and position information. Passive tags required power from a reader to operate, so the transmission of information from the tag was limited to those times when it was in close proximity to a reader. The integrated power source of an active tag allows it to broadcast continuously, or at regular intervals, allowing the overall system to maintain real-time information about the identity and location of resources.

The final piece of technology needed to realize an IPS based on RFID requires the use of enhanced readers and a central software application to compute position. There are several competing technologies currently in the marketplace, but all require that multiple readers be deployed in the area to be monitored and that the data from the monitors is routed to some central point for processing. The systems range from simple configurations that establish location with room-level granularity by placing a reader in each room of the area to be monitored, to high precision systems that use triangulation of transmitter power to establish location relative to three or more readers.

### **c. Issues in Healthcare**

The use of RFID in healthcare presents a number of critical issues, some of them unique to healthcare, some basic to the technology.

*Privacy*—There is already public concern over RFID due to its perceived ability to facilitate passive monitoring of the actions of people. Since this concern has already arisen during the early application of RFID, which was principally restricted to objects in the retail supply chain, it can be expected that the level of controversy will increase once the tagging of people becomes commonplace. Placing this system into a healthcare environment already rife with privacy issues will require careful attention to the perceived impact on those being monitored be they patients, staff or physicians.

*Radiation*—Another issue with a significant social component is the health consequences of exposure to RF radiation from readers and tags. Though vendors manufacture their devices in accordance with relevant standards for human safety in electronic fields, the integration of these devices into a system incorporating multiple readers and transceivers affixed directly to humans will necessitate a system-level evaluation of safety.

*Electromagnetic Interference*—RFID systems will emit electromagnetic, infrared, or other types of radiation, depending on the underlying technology used. The healthcare environment is already full of safety critical devices that are sensitive to radiation at various frequencies. Care must be taken to insure that RFID systems can co-exist with current and future systems for diagnosis, monitoring or treatment.

*Environmental Hazards to Tags*—The healthcare industry presents a unique challenge to the physical integrity of RFID tags because of its pervasive infection control measures. Tags affixed to clothing, even temporarily, will almost certainly be sent to the laundry from time to time. Tags affixed to some supplies and equipment will be subjected to sterilization by thermal or chemical means. Because the original target market for RFID tags

was the relatively benign retail supply chain, most COTS tags will not stand up to harsh environments. Thus care must be taken to select tags appropriate to the environment of any item they will be affixed to.

*Identifying Individuals by Proximity*—Care must be taken in using a non-line-of-sight technology like RFID to identify individuals. If two tagged persons or objects are in close proximity to one another, even the most accurate RFID technology will have difficulty differentiating the two. Using such a system to make safety critical judgments of the identity of individuals is not advised.

*Buying Technology vs. Integrated Systems*—Care must be taken in vendor selection to insure that the purchased system architecture is sufficiently open to allow integration with existing and future systems. A solution to the location problem that doesn't allow location data to be fed to mission-specific applications is of little utility.

#### **d. Technologies**

In this section we survey technologies that can be used to create an indoor positioning system that will solve the location problem.

*Magnetic stripe cards* are a low-cost, reliable technology that can be used to identify reliably identify the card. If the physical locations of the card readers are known then the location and identity of a card can be recorded at the time it is swiped.

Costs for cards and readers are quite low due to the volume of cards and readers used in the general marketplace. Cards are rewritable, and a small amount of information beyond the card identifier can be stored on the card. No radiation of any kind is emitted by cards or readers.

Cards are easily lost, and without the added expense of imprinting (an ink label or photograph) cards can be accidentally interchanged thereby producing erroneous identifying information for the bearer. Magnetic stripe cards are easily replicated and should not be used in applications where fraudulent use would be harmful. Since cards must come in contact with their readers there are periodic maintenance issues for the cards and readers (card readers contain parts that wear with use; card stripes need to be cleaned and inspected for physical damage periodically). Cards are typically not tolerant of extreme heat or exposure to various chemical agents. The data stored on the card is subject to disruption by magnetic fields.

*Bar coding systems* provide a system similar to magnetic stripe cards for the identification of items displaying a bar code. However the reading of bar code labels requires only a line of sight between bar code and reader, rather than the physical contact required for the swiping of magnetic stripe cards. This has the advantage of reduced maintenance costs and more flexibility in the placement of bar codes. Bar codes can be printed anywhere a line of sight is available to the reader, whereas magnetic stripe cards must conform to standards for the placement of the magnetic stripe on a card of fixed dimensions that is free to be physically swiped through a reader.

Costs are relatively low. Bar codes are read-only, with their value set at the time the label is created. Bar code readers emit visible light. In typical applications this light is harmless to humans, but some care must be taken in its use. Labels can be printed using technologies that are impervious to any sterilization procedure.

Though no physical contact is required between bar code and reader, a line of sight must be available for reading. Also, most applications require human intervention to properly align the label and/or reader to get an accurate recording of data. The duplication of a bar code is a trivial exercise and thus bar coding schemes should not be used in applications where fraudulent use would be harmful.

*True RFID systems* using either active or passive tags provide a system similar to magnetic stripe cards or bar coding systems without the need for physical contact or even a line of sight between tag and reader. Costs are significantly higher than magnetic stripes or bar codes, an active RF signal will be present in the environment being monitored and sterilization tolerance requires special care in tag design. The choice of tag and reader will dictate the size and accuracy of the read zone; low cost, low power systems will tend to have read zones that allow sampling of identification and location data only in close proximity to the reader, not unlike magnetic stripe and bar coding systems. RFID tags are difficult, but not impossible to duplicate. Some application-specific COTS RFID tags intended for high security applications include features that make them virtually impossible to duplicate.

*Infrared RFID systems* use infrared (IR) light (like a television remote control) rather than radio frequencies to communicate between tags and readers. The use of IR means that tag reading can be done with casually aligned tags and readers any time the tag's emitted signal can reach a reader. (Contrast this with bar code reading which requires the reader to be carefully positioned with respect to the tag being read.) An advantage of IR over true RFID is that IR signals are more easily blocked, thus a signal emitted in a room surrounded by opaque partitions will not usually be read in an adjacent room. A corresponding disadvantage is that IR systems are line of sight, therefore any tags obstructed by clothing, equipment dust covers or supply packaging cannot be read.

*Hybrid RFID/IR systems* exist that give the non-line of sight performance of an RF system coupled with the ability to easily isolate IR signals. Such systems typically use the RF signal to identify the tag, and when more than one reader receives a tag's signal IR readers will attempt to resolve the ambiguity. However this tie-breaking ability will only work when there is a line of sight from tag to IR reader.

*Wi-Fi RFID systems* attempt to exploit existing wireless network infrastructure (*i.e.*, Wi-Fi access points) for use as RFID readers. Tags must communicate using one of the 802.11 wireless protocols, and the circuitry is therefore relatively complex (compared to traditional RFID). Using Wi-Fi RFID tags, an indoor positioning system (IPS) can be constructed using existing Wi-Fi network infrastructure. However, determination of location to a fine degree of precision may require significantly more Wi-Fi access points than were provisioned for the original 802.11 wireless network.

#### **e. Standards Activity**

In this section we list national and international standards for the manufacture and operation of RF-based RFID systems. Though national regulators responsible for control of the RF spectrum may require conformance with specific air interface standards, participation in other standards is strictly voluntary and conformance will vary by vendor.

Standards are promulgated by various international standards bodies including ANSI, ISO, IEC and INCITS, in addition to a pair of industry specific standards bodies. EPCglobal (<http://www.epcglobalinc.org>) is a standards setting body for electronic product codes (EPC) and RFID. The Association for Automatic Identification and

Mobility (AIM, <http://www.aimglobal.org>) is a trade group that provides standards advisory for bar coding and RFID systems.

ISO/IEC 18001 is the air interface standards for RFID systems transmitting or receiving in the RF spectrum. The standard has six parts addressing generic properties of transceivers as well as properties of transceivers operating in specific frequency bands.

ISO/IEC 15961, 15962, 15418 and 15434 prescribe command and data interchange formats applicable to RFID systems.

ISO/IEC 15963 and 15459 create systems for insuring the uniqueness of tag identifiers.

ANSI 256-2001 establishes a standard for creating interoperable RFID systems within the United States.

EPC Version 1.0 Specifications provides a minimum requirement for creating an “EPCglobal Network” compliant RFID system based on the Electronic Product Code (EPC) standard. It is intended for use in global supply chain monitoring.

EPC Generation 2 Enhances the EPC Version 1.0 standard with the addition of UHF transceivers. This standard has received a great deal of attention because it has been mandated for use at the case and pallet level by Wal-Mart.

INCITS 371 is being developed by INCITS working group T20 for Real Time Locating Systems. It is intended to cover a UHF and VHF air interface protocol and an API for creating standards compliant real-time locating systems (*i.e.*, a generalization of IPS).

#### **f. Proposed Selection Criteria**

We close our technology discussion with nine key criteria for evaluating technologies for use in creating an IPS for a healthcare environment.

##### *i. What accuracy of location determination will the system support?*

The key here is not to find the most accurate system possible, but rather to find a system that provides the accuracy level required by the application envisioned. It would be possible to build an employee attendance system using an IPS, but there are more cost effective means (*e.g.*, punch clocks) to accomplish the same end. Likewise, it may be possible to drive a real-time workflow engine using bar code scanners, but the burden of frequent scanning would make the system unwieldy. Even within RFID-based indoor positioning systems, the accuracy of location reporting varies greatly, and the technology must be closely evaluated to insure that the accuracy provided is necessary and sufficient for the planned use.

##### *ii. How accurate is the location reporting under realistic operating conditions?*

Ideally (from the point of view of the technology) an RFID-based indoor positioning system would operate in an environment completely free of materials capable of reflecting or absorbing RF emissions,

with readers uniformly distributed across the environment being monitored. In practice the healthcare workplace is full of objects both reflective and absorptive, and such objects are added, removed or moved over time. The configuration and estimated accuracy of the system should be determined with careful attention to the floor plan, furnishings and workflow of the environment being monitored. Likewise, the physical placement of readers will be dictated in large part by pragmatic concerns like the availability of power, network connectivity and an appropriate physical mounting point. The impact of non-optimal placement of readers must be understood at design-time. Since they depend on the ability of typical partitions to completely block the passage of light, IR systems can be especially tricky to configure for open workspaces, or workspaces separated by glass partitions.

iii. *How passive is the technology?*

Does the technology provide the required accuracy without human intervention, or does it rely on humans to mark sentinel events by swiping, passing through a specific portal, or button-pressing? Is the level of human intervention required practical for the application? If not, what will be the impact on the usefulness of the overall system if the location data is spotty or absent for intolerant users?

iv. *What density of readers will be required to support the required accuracy?*

Understand the connection between reader density and placement and the level of accuracy that can be expected. The fact that 802.11-based indoor positioning systems can make use of your existing Wi-Fi infrastructure is appealing, but your network infrastructure was not engineered for this purpose. Achieving the desired level of accuracy will require analysis of the existing network to determine if it will be necessary to add or move access points.

v. *Do available tags satisfy the unique requirements of the operating environment?*

Consider the physical tags themselves. Can they be mounted to the items you intend to track without interfering with the use of those items? Can they tolerate the same physical abuse, cleaning, heating or neglect that the tagged item must? Have any electromagnetic emissions been certified for compatibility in the intended environment?

vi. *How can the technology be integrated with existing and planned systems?*

Is the technology delivered as a single closed system, or does it have an open architecture that can be integrated with existing and planned systems? Can the technology share location data in real time? Will the technology respond to queries, or is that functionality to be implemented at higher levels?

vii. *What maintenance will be required to support the system under normal use?*

Maintenance tasks come in several different varieties, including: (1) Physical maintenance: What are the service intervals of components subject to physical use? What are the labor and parts costs for the maintenance of components requiring batteries? (2) Configuration maintenance: Once the system is established, how will the dynamism of the environment impact its functioning and accuracy? Will simply moving cabinets, equipment or partitions have an impact; if so, what re-design and re-

deployment activities will be required to restore maximum functioning and accuracy? What are the long range plans for remodeling and new construction and how will they impact the system? (3) Software maintenance: How will the technology be integrated with existing and future IT systems? How will faults in operation, including failures, inaccurate reports, and safety, security and privacy issues be resolved?

viii. *What is the total lifecycle cost to design, build, train staff to use, and maintain the technology?*

Up-front equipment costs are typically a small fraction of the lifecycle costs of any complex system. Be careful to account for the cost of initial analysis and design activities, periodic review of performance and re-engineering as necessary, training of staff, management of tags, software maintenance and upgrades, *etc.*

ix. *What is the projected return on investment?*

What qualitative and financial benefits are anticipated from the system? How are they quantified? At what rate will these benefits be realized?

### **3. Applications**

In this section we present a survey of healthcare RFID projects reported in the trade press and press releases. This is not intended as an all-encompassing review of RFID deployment at hospitals across the country, but rather as a survey of published accounts of such projects. These projects fall into four broad categories including patient identification and electronic medical records, automated verification, infant security and workflow monitoring and management.

#### **a. Patient Identification**

Patient identification projects focus on accurate identification of patients throughout the perioperative process. RFID-based patient identification projects are a natural extension of bar-code patient identification systems that offer the ability to dynamically update stored information that travels with the patient.

In June of 2004 CIMIT and Precision Dynamics Corporation (PDC) announced a project to test the efficacy of RFID wristband technology to improve patient safety, reduce errors and improve workflow. PDC's technology uses a passive RFID tag embedded in a disposable plastic patient ID bracelet. Tags and readers communicate at 13.56 MHz. Readers are available in various configurations supporting a read range of 5-8 inches [1,2]. A similar project is underway at Ohio State University Medical Center.

The Navy's Tactical Medical Coordination System (TacMedCS) uses RFID dog tags to identify casualties on the battlefield. The tags store identifying information and injury and treatment information as soldiers move through the treatment process. This system is meant to replace a paper "triage tag" that is currently used to record the medical record of a casualty case as it moves from battlefield through treatment [3,4,5].

The SurgiChip Surgical Marker [6] is an RFID-based system to support JCAHO's "Universal Protocol." The tag is affixed to the proper surgical site with adhesive and remains in place until treatment is complete. The system is intended to aid hospitals in reducing wrong-site, wrong-procedure and wrong-patient surgeries.





**Figure 4: Personal Information Carrier**

The Army is developing prototype electronic dog tags like that shown in Figure 4. The current generation Personal Information Carrier (PIC) shown in the figure is a high capacity flash memory in a ruggedized housing. It is capable of storing large volumes of medical or operational data, accessible through an edge connector. This device transmits and receives data through a physical connection, but the next generation PIC, known as the Electronic Information Carrier (EIC), will replace the edge connector with an RFID interface for contactless access to stored data.

#### **b. Automated Verification**

Automated verification applications focus on reducing or eliminating medical errors by detecting the people and/or objects that will be involved in an activity and automatically

verifying that those participants are valid.

In June of 2004 Massachusetts General Hospital (MGH) reported preliminary experience from a trial using 13.56 MHz passive RFID technology to reduce errors in blood transfusion. Patients wore identification bracelets with embedded RFID tags, and blood containers were outfitted with their own embedded RFID tags. A short-range RFID reader is used at the bedside to verify that patient tag and blood container tag are compatible. The researchers envision the use of the system in operating rooms where passive monitoring of the identity of the patient and blood products present in the room could prevent serious errors. A similar project is being conducted at the Georgetown University Hospital with an emphasis on comparing the relative efficiencies of bar codes and RFID [7,8].

In February of 2004 The US Food and Drug Administration (FDA) advocated the use of RFID to establish the provenance of pharmaceuticals [9,10]. FDA envisions a system of “mass serialization” in which each unit of product would be given a unique serial number, stored in an RFID tag attached to that unit and carried throughout its lifecycle. The serial number would provide access to a record of manufacturing data, shipment points, and current location throughout the pharmaceutical supply chain from manufacturer up to (but not necessarily including) end user. Such a system is believed to offer a number of benefits including protection against counterfeiting, simplified inventory management, rapid, targeted recalls, prevention of diversion, and confirmation of correct dispensing of prescriptions. Deployment is expected to take at least four years due to necessary regulatory, standards, and system engineering issues. In November of 2004 FDA released regulatory guidance meant to eliminate obstacles to the use of RFID tags in pharmaceuticals manufacturing and distribution [11,12].

Colder Products Company is marketing “Smart Coupling Technology” that uses RFID tags and readers in fluid connectors. The system reports the time, date and location when couplers are joined, and can be used to insure that the fluid source and sink are not being connected in error [13,14,15].

In February 2005 a task force investigating the 1999 scandal at UC Irvine involving the diversion of human tissues recommended the use of RFID tags to prevent unauthorized distribution of cadavers. Tagging would



allow auditors to use readers to quickly verify identity of tissue samples and verify a facility's inventory [16,17,18].

### **c. Infant Security**

Doctors Hospital of Dallas uses an RFID tagging system from Xmark Systems for infant protection. Babies are tagged with a tamper-proof anklet at birth that periodically sends out a signal indicating their presence. The absence of a signal in the obstetrics unit, or an unauthorized attempt to pass an infant through a door exiting the unit will raise an alarm. Mothers are also tagged with an active bracelet tag that acts as a reader for infant anklet tags. This two tags working together prevent incorrect matching of mothers to infants during the hospital stay for activities such as breast feeding. According to Xmark the system is in use at more than 400 hospitals across the United States [19,20].

### **d. Workflow and Equipment Monitoring and Management**

Workflow monitoring and management and equipment monitoring are by far the most broadly deployed RFID systems in the healthcare arena. The degree of integration of the location data with workflow systems varies from simple equipment location systems to workflow-centric systems that attempt to detect and trigger events using RFID technology.

St. Vincent's Hospital, Birmingham, AL, working with Awarix, has deployed an RFID enabled workflow engine to drive clinical processes. A key feature of this system is the Awarix Patient Care Communication Board that replaces the traditional departmental white board with real-time status of tasks and location of resources based on sources that include traditional RFID [21]. Similar work is also being deployed at Hannibal, Missouri, Regional Hospital [22]. Albert Einstein Medical Center in Philadelphia is using a system from Versus Technology to track the flow of staff in their emergency medicine department [23,24].

In October of 2004 Massachusetts General Hospital (MGH) received \$1.5 Million of NIH funding for an 18-month project to use a Radianse IPS to measure patient care processes. Surgical patients will receive Radianse tags as they enter the treatment process, and be tracked throughout their stay. A complete record of location, service times and wait times will be accumulated for surgical patients, and used to evaluate and improve the quality of care processes at MGH [25].

The Hospital of the University of Pennsylvania (HUP) is using a Radianse IPS for asset tracking across four buildings of its medical complex. HUP is using a combination of room and zone-level location precision to track real-time location of medical equipment, devices and accessories. The goals of the project are to improve equipment utilization, reduce losses, and increase clinician satisfaction and productivity [26,27]. Similar projects are under way at Beth Israel Deaconess Medical Center, Brigham and Women's Hospital, Vanderbilt Children's Hospital, and Bon Secours Richmond Health System. The Navy's Fleet Hospital Support Service is using RFID for asset tracking in the deployment of field hospitals [28,29].

Vanderbilt Children's Hospital has equipped ICU refrigerators with RFID chips capable of monitoring temperatures in real-time [30]. A similar project for RFID monitoring of the supply chain temperature of drug eluting stents was reported by Mercy Hospital at the 2nd Annual RFID, Tracking & Barcoding for Hospitals conference in January 2005 [31].

North Bronx Healthcare Network reports [32] using RFID tags on patients to facilitate rapid, accurate identification and improvement of workflow. Patients are tagged upon admittance, and caregivers use wireless tablet PC's to read the tags and access patient medical records and clinical records in real-time.

The Washington Hospital Center, Washington, DC has contracted with Parco Wireless for the installation of an ultra-wideband RFID system throughout their Emergency and Ambulatory Care Departments. The system provides sub-foot accuracy in location reporting. Equipment and staff will be tagged with active tags that can be read from up to 600 feet away. Data recorded will be used as input for various unspecified systems [33,34,35].

#### **4. Vendor Review**

In this section we review prominent vendors in the areas of RFID and IPS for healthcare applications. This is by no means an exhaustive list as new consultancies are created almost daily to aid in the design and deployment of RFID and IPS solutions for healthcare. The emphasis here is on firms with an established presence and proven track record.

##### **e. Low-Level RFID Technology**

The dominant vendors in traditional RFID tag and label technology are Philips Electronics and Texas Instruments. Philips I-CODE labels and UCODE-based tags and labels are basic building blocks for creating passive tag RFID solutions. Texas Instruments Tag-IT system provides competing products for similar applications.

Detailed technical information is available online:

I-CODE: <http://www.semiconductors.philips.com/markets/identification/products/icode/index.html>

UCODE: <http://www.semiconductors.philips.com/markets/identification/products/ucode/index.html>

Tag-IT: [http://www.ti.com/tiris/default.htm?DCMP=TIHomeTracking&HQS=Other+OT+home\\_tirfid](http://www.ti.com/tiris/default.htm?DCMP=TIHomeTracking&HQS=Other+OT+home_tirfid)

##### **f. Healthcare Oriented RFID and IPS**

At present the market for healthcare oriented IPS is fragmented, with no technology having yet become dominant, and most firms having very narrow product offerings. A number of prominent players exist, all emphasizing slightly different technology or systems.

- Ekahau—Ekahau, Inc., focuses on providing software and tags to create an IPS from existing Wi-Fi network infrastructure. A system is created using Ekahau's patented site calibration tools that compute a signal strength map during a pre-installation site survey. This map is used in real-time by Ekahau's Positioning Engine (EPE) software to locate any Wi-Fi enabled device. The positioning engine works with industry standard Wi-Fi devices, or application specific tags can be purchased directly from Ekahau. Claimed accuracy of position reporting is one meter. Ekahau provides tools and a documented application programming interface for integrating EPE location information into custom middleware or end-user applications.

Additional information is available online at <http://www.ekahau.com>.

- Exavera Technologies—Exavera sells specialized network hardware for creating integrated Wi-Fi, RFID and voice over IP networks. They sell specialized passive and active RFID tags intended for use in healthcare settings. Passive tags have read ranges of up to 45 feet, active tags up to 90 feet. Integration with custom middleware or end-user applications is through XML messaging.

Additional information is available online at <http://www.exavera.com>.

- Mobile Aspects—Mobile Aspects sells a collection of RFID-enabled systems focused on medical supply inventory management, patient tracking, equipment tracking and intelligent supply stations for anesthesia and patient bedside applications. The inventory and supply applications implement automated inventory control using passive RFID. Supply applications also include intelligence for notifying the user regarding potential drug interactions. Patient and equipment tracking are based on active RFID tags.

Additional information is available online at <http://www.mobileaspects.com>.

- PanGo Networks—PanGo markets software for creating IPS from Wi-Fi network infrastructure. The PanOS platform provides the low-level functionality for determining location of Wi-Fi enable devices in real-time. PanOS is integrated with a number of generic IPS applications for asset tracking and display, and provides an application programming interface for integrating with custom middleware or end-user applications.

Additional information is available online at <http://www.pangonetworks.com>.

- Parco Merged Media—Parco markets an ultra-wideband (UWB) RFID/IPS system. Parco claims their UWB technology is unique among healthcare IPS providers and provides immunity from eavesdropping, interference and jamming that 802.11-based networks cannot provide. Parco provides a historical record database, and tools and a documented application programming interface for integrating location information into custom middleware or end-user applications.

Additional information is available online at <http://www.parcowireless.com>.

- Radianse—Radianse provides technology for creating IPS systems based on combined RF and infrared (IR) location technologies. Tags emit both RF and IR signals which are detected by readers which attempt to assign the location of a tag to the nearest reader based on received RF and IR signals. Since many building materials pass RF signals but not IR, pure RF can be used for coarse position determination, and where needed, IR signals can refine position information to a particular partitioned space. Radianse sells wristband tags for tracking patients, and generic tags for all other applications. Access point hardware is proprietary, but can be connected into existing wired LAN infrastructure. Location data can be accessed through Web, database or XML interfaces for integration into custom middleware or end-user applications.

Additional information is available online at <http://www.radianse.com>.

- UbiSense—UbiSense markets an ultra-wideband (UWB) IPS system with location accuracy of six inches. Higher accuracy is made possible by the UWB signals which are less prone to multipath distortion than traditional RF signals. UbiSense claims their systems require a lower reader density than traditional RF-based systems due to the use of more elaborate position resolution algorithms. The vendor provides tools and a documented application programming interface for integrating location information into custom middleware or end-user applications.

Additional information is available online at <http://www.ubisense.net>.

#### **g. IPS Enabled Workflow Solutions**

A number of firms focus on the application of workflow concepts to hospital environments. Three prominent companies that provide IPS enabled workflow engines are Agility Healthcare Solutions (<http://www.trenstar.com/agility/>), Awarix (<http://www.awarix.com>) and PeriOptimum (<http://www.perioptimum.com>).

#### **h. Healthcare Oriented RFID Materiel**

Finally, many vendors of medical supplies and equipment are beginning to integrate RFID or IPS technology into their products. A sampling of these items is as follows:

- ClearCount Medical Solutions—ClearCount is a start-up that intends to market surgical sponges with integrated RFID tags. These sponges will replace the traditional manual inventorying of sponges during surgery with a wand procedure that will tell definitively whether or not tagged sponges remain in the surgical site.
- Colder Products—Smart Coupling technology integrates RFID tags and readers into fluid couplings to verify the appropriateness of connections in real-time, and to create a historical record of coupling and uncoupling activities.

Additional information is available online at <http://www.colder.com>.

- Precision Dynamics—Precision Dynamics has added wristbands with embedded passive RFID tags to its line of patient identification wristbands.

Additional information is available online at <http://www.pdcorp.com>.

- SRI/Surgical Express—This hospital supply company has sewn passive RFID tags into surgical gowns and drapes to expedite processing of soiled linens.
- SurgiChip—SurgiChip is an FDA approved RFID tag specifically marketed for surgical site identification. The system is intended for use in the JCAHO Universal Protocol program to help prevent wrong-site, wrong-procedure and wrong-patient surgery.

Additional information is available online at <http://www.surgichip.com>.

- VeriChip—The VeriChip RFID tag is a miniaturized RFID transceiver specifically designed to be implanted under the skin. The vendor claims that it provides secure, tamperproof identification of individuals for medical, financial and security applications.

Additional information is available online at <http://www.4verichip.com>.

## 5. Literature Review

In this section we review published literature related to the application of RFID and Indoor Positioning Systems in health care settings. Though there was no directly relevant material in archival medical informatics journals, there is a wealth of information available online. There are also significant technology- and application-oriented articles to be found in the engineering research literature.

### a. Thomson ISI Indexed Journals

Thomson ISI indexes eighteen English language journals related to Medical Informatics. They are:

1. *Journal of the American Medical Informatics Association*, published for the American Medical and Informatics Association by Hanley & Belfus, Inc.
2. *Journal of Biomedical Informatics*, published by Elsevier, Inc.
3. *International Journal of Medical Informatics*, published for the International Medical Informatics Association (IMIA) and the European Federation of Medical Informatics (EFMI) by Elsevier, Inc.
4. *Methods of Information in Medicine*, published by Schattauer Publishers.
5. *Artificial Intelligence in Medicine*, published by Elsevier, Inc.
6. *Medical Decision Making*, published by Sage Publications, Inc.
7. *Journal of Evaluation in Clinical Practice*, published by Blackwell Publishing.
8. *IEEE Engineering in Medicine and Biology Magazine*, published by IEEE.
9. *IEEE Transactions on Information Technology in Biomedicine*, published by IEEE.
10. *Computer Methods and Programs in Biomedicine*, published by Elsevier, Inc.
11. *International Journal of Technology Assessment in Health Care*, published by Cambridge Journals.
12. *CIN-Computers Informatics Nursing*, published by Lippincott, Williams and Wilkins.
13. *Journal of Medical Internet Research*, (eJournal) published by the Centre for Global eHealth Innovation.

14. *Medical and Biological Engineering and Computing*, journal of the International Federation for Medical and Biological Engineering, published by IEE.
15. *Medical Informatics and the Internet in Medicine*, published by Taylor and Francis.
16. *Journal of Cancer Education*, published by Lawrence Erlbaum Associates.
17. *Statistics in Medicine*, published by Wiley.
18. *Statistical Methods in Medical Research*, published by Hodder Arnold Journals.

A search of all eighteen journals using the keywords “RFID,” “IPS,” and “Indoor Positioning” returned one relevant hit from the *Journal of Medical Internet Research*. In this article the term “RFID” received passing mention in an opinion item on new technologies.

### **b. Web Sources and Trade Publications**

Articles on RFID and indoor positioning systems are much more plentiful in web sources and trade publications. These include application reports from the popular press [3,16,18,23], vendor press releases [1,5,25,27], summaries of projects disseminated by the participating hospitals, vendors or government agencies [4,8,10,11,12,13,19,22,26, 28,29,30,31,32,35], and technical data from vendor websites [6,14,20,21,24,34]. A number of online information sources exist solely to track technology trends related to RFID and indoor positioning, including *RFID Journal* [36] and *IDTechEx* [37].

### **c. Engineering/Computer Science Literature**

In the engineering and computer science literature the areas of “pervasive computing,” “ubiquitous computing,” and “location aware computing” are relevant to RFID and indoor positioning technologies. This literature covers a wide range of topics, from low-level technology through attempts to infer meaning from sensor data up to applications. Though much of the writing is too technology-oriented or abstract to be directly applicable to our presentation, some ideas do resonate in a health care context.

Accurate association of persons with objects they have used is a subtle problem, given current technology. The presence of person and object in close proximity is at best a very coarse indicator, as people may come into close proximity with many objects they don’t actually use. A more accurate, though cumbersome approach involves tagging objects with sensors containing an accelerometer, clock and local memory [44,45]. If an object is jostled while a person is in close proximity, it may be reasonable to assume that the person has made some use of that object. Fishkin, Jiang, Philipose and Roy attempt to solve the problem using only information that can be inferred from the RF signals returned by objects [43]. They demonstrate that despite very limited information on signal strength and quality they can reasonably infer interactions between people and objects in home-use settings.

Their work is interesting because it demonstrates extraction of interesting high-level properties using off-the-shelf technologies in realistic settings. Their treatment of various factors that impacted signal strength (and their inferred results) provides a good introduction to various real-world problems with accurate RFID sensing. In a health-care setting their ideas could provide a starting point for determining interaction between tagged entities at a more accurate level than simple co-location.

Inferring high-level behavior from sensor data and using that data to create predictive models of future behavior is studied in the context of GPS by Patterson, Liao, Fox and Kautz [38]. Their motivating example was the Activity Compass [39], a tool that helps guide a cognitively impaired person safely through their normal routines, noting when they have deviated from their predicted path. The tool uses GIS map data, GPS position samples, and historic travel paths to deduce the routes and modes of transportation employed over the travel paths. The inference is done automatically using a novel combination of Bayes filters, graph-constrained particle filtering and Expectation-Maximization [40,41,42].

One could imagine such a system being employed directly in a health care setting. For example, floor maps and IPS data could be used to infer work patterns and detect deviations from established work patterns. Or occasional visitors to the hospital could be given a tool very much like the Activity Compass to guide them during their visit.

The health care related application of pervasive computing receiving the most attention in the computer science literature is the use of ubiquitous computing to aid the cognitively impaired. Researchers at Intel Research have been focusing their efforts on understanding how computing can best help (or better yet, not hinder) the cognitively impaired to carry out their daily routines [46,47]. Their work has focused on the study of Alzheimer's patients, and an attempt to define how much context a computer system must be aware of to keep pace with a human's need to understand context.

Though presented in the context of aids for the cognitively impaired, the basic idea of maintaining consistent context between humans and computers is not new. In the 1990's the aviation safety community determined that *mode confusion* played an important role in a number of aviation accidents and near misses [48,49,50]. In mode confusion a pilot's notion of the mode of his instruments and the actual operating mode of those instruments differs, leading to situations where, for example, the aircraft's instruments are aware that the airplane is very close to the ground but a pilot interpreting altitude data in the wrong units believes the plane to be flying at a safe altitude. As ubiquitous sensors and systems that reason based on sensor data play a larger role in health care, the need to maintain consistent notions of context between man and machine will become ever more important.

At a conceptual level, the Centre for Pervasive Computing at the University of Aarhus, Denmark, is pioneering an activity-centered perspective for modeling computing systems in the health care setting [51]. They caution that as computing becomes ubiquitous in the health care setting, it must be sensitive to the fact that health care work is different from typical office work. The need for mobility, ad-hoc collaborations, interruptions, and high degree of communication make this type of work ill-suited for traditional modes of automation, *i.e.*, application-oriented systems based on complex interactions between a single system and human, and workflow-oriented systems based on rigid allocations of work by a system.

## 6. Summary

In summary, the development, marketing and deployment of RFID and IPS for healthcare applications is in its early stages. The market is divided among a number of competing technologies based on non-standard and incompatible technologies. There have been many pilots, though the reporting on these projects provides little quantitative information on the performance or earned value of the systems.

A major point that hampers understanding in the field is a lack of consistent terminology. Though there is much talk of deploying “RFID” in healthcare, most references to such systems are actually meant to describe real-time indoor positioning systems based on relatively high-power RF or IR transmitters and receivers. The application of true RFID technology is limited to simple identification tasks (wrist bands with embedded RFID tags) and supply chain management.

Most deployments seem to be motivated by a desire to experiment with new technology, or to reduce the burden on hospital staff by making it easier to track the location of assets. Several vendors provide quantitative arguments based on standard estimates of the cost of medical errors. It seems unlikely any such system alone would provide an overall reduction in the rate of medical errors necessary to actually pay for the system’s deployment and ongoing operation. Other financial arguments are made at a qualitative level by asserting increased equipment utilization as a result of location technology, and corresponding reductions in capital costs for equipment. At present a solid business case and financial model is lacking for decision making in advance of deployment and assessment following implementation.

Potential valuable contributions to the field include (1) independent, quantitative reporting of the accuracy and value of systems deployed in real work environments, (2) an archival quality reporting of such results, and (3) uniform standards for terminology and technology to facilitate discussion, comparison and interoperable implementations.



## References

- [1] "Precision Dynamics Corporation (PDC) joins as newest member of CIMIT's industry liaison program," CIMIT press release. [http://www.cimit.org/news/pdc\\_joins\\_ilp.html](http://www.cimit.org/news/pdc_joins_ilp.html)
- [2] Precision Dynamics Corporation corporate website. <http://www.pdcorp.com/rfid/>
- [3] Munsey, C., "New tracking system relays casualty data from battlefield," *Navy Times*, March 1, 2004.
- [4] International Quality and Productivity Center corporate website. <http://www.iqpc.com/cgi-bin/templates/document.html?topic=483&event=5992&document=47445>
- [5] "Virtual medical system beams Navy into 21<sup>st</sup> century," Pacific Northwest National Laboratory press release, May 31, 2000. <http://www.pnl.gov/news/2000/00-26.htm>
- [6] Surgichip corporate website. <http://www.surgichip.com>
- [7] IDTechEx case study, February 26, 2005. <http://www.idtechex.com/knowledgebase/en/casestudy.asp?freefromsection=5>
- [8] "Georgetown University Hospital currently piloting TransfuseRFID," AMTSystems corporate website. <http://www.amtsystems.com/Health/templates/news.cfm?id=56>
- [9] "FDA endorses RFID technology," *RFID Journal*, February 18, 2004. <http://www.rfidjournal.com/article/view/801/1/1/>
- [10] FDA, *Combating Counterfeit Drugs: A Report of the Food and Drug Administration*, February 2004. [http://www.fda.gov/oc/initiatives/counterfeit/report02\\_04.pdf](http://www.fda.gov/oc/initiatives/counterfeit/report02_04.pdf)
- [11] Collins, J., "FDA clears way for RFID tagging," *RFID Journal*, November 15, 2004. <http://www.rfidjournal.com/article/articleview/1238/1/1/>
- [12] FDA, "Radiofrequency identification feasibility studies and pilot programs for drugs," *Compliance Policy Guides*, Sec. 400.210, November 2004. [http://www.fda.gov/oc/initiatives/counterfeit/rfid\\_cpg.html](http://www.fda.gov/oc/initiatives/counterfeit/rfid_cpg.html)
- [13] Garber, R., "Is this the right drug? RFID tag tells," *Medical Design*, April 1, 2004. <http://www.medicaldesign.com/articles/ID/12200>
- [14] Colder Products Company corporate website. [http://www.colder.com/asp\\_Main/SmartTechnologyProducts.asp](http://www.colder.com/asp_Main/SmartTechnologyProducts.asp)
- [15] Colder Products Company corporate website. [http://www.colder.com/asp\\_SmartCoupling/SmartCouplingTechnologyHomePage.asp](http://www.colder.com/asp_SmartCoupling/SmartCouplingTechnologyHomePage.asp)
- [16] "RFID tags proposed to halt blackmarket cadaver trade," Technovelgy.com website, February 10, 2005. <http://www.technovelgy.com/ct/Science-Fiction-News.asp?NewsNum=331>
- [17] "UC may use barcodes, RFID for cadaver security," SiliconValley.com website, February 4, 2005. <http://www.siliconvalley.com/mld/siliconvalley/10818955.htm>
- [18] Associated Press, "UCLA considers cadaver barcodes," February 4, 2005. <http://www.cnn.com/2005/EDUCATION/02/04/cadaver.barcodes.ap/>
- [19] "The infant and mother matching system," *Smart Labels Analyst*, Issue 41, June 2004. <http://www.idtechex.com/documents/en/sla.asp?documentid=172>
- [20] Xmark corporate website. [http://www.xmarksystems.com/product\\_pages/hugs\\_advantages.html](http://www.xmarksystems.com/product_pages/hugs_advantages.html)
- [21] Awarix corporate website. <http://www.awarix.com/>
- [22] Goedert, J., "Hospital realizes benefit of patient tracking," *Mobile Health Data*, December 1, 2004. <http://www.mobilehealthdata.com/article.cfm?articleid=1156>
- [23] Fitzgerald, T., "Cart 54, where are you? The tracking system knows," *New York Times*, October 30, 2003. <http://www.radianse.com/download/news-nyt-oct2003.pdf>
- [24] Versus Technology corporate website. <http://versustech.com/>

- [25] "NIH awards Radianse \$1.5 million to test the impact of real-time location on healthcare efficiency, safety," Radianse corporate press release, October 26, 2004. <http://www.radianse.com/press-nih-102604.html>
- [26] "The Hospital of the University of Pennsylvania installs Radianse indoor positioning system," *RFIDNews*, January 12, 2005. <http://www.rfidnews.org/news/2005/01/12/the-hospital-of-the-university-of-pennsylvania-installs-radianse-indoor-positioning-solution/>
- [27] "The Hospital of the University of Pennsylvania installs Radianse indoor positioning system solution across PeriOperative Services," Radianse corporate press release, January 11, 2005. <http://www.radianse.com/press-HUP-011105.htm>
- [28] DoD—Supply Chain Integration Initial Implementation Lessons, May 13, 2004. <http://www.dodait.com/conf/ipt/0504/Day2/Lessons%20Final%20051204.ppt>
- [29] International Quality and Productivity Center corporate website. <http://www.iqpc.com/cgi-bin/templates/document.html?event=5992&topic=535&document=47446&>
- [30] Krizner, K., "Technology investments can boost patient safety, monitor trends," *Managed Healthcare Executive*, July 1, 2004. <http://www.managedhealthcareexecutive.com/mhe/article/articleDetail.jsp?id=105392>
- [31] International Quality and Productivity Center corporate website. <http://www.iqpc.com/cgi-bin/templates/document.html?event=5992&topic=535&document=47446&>
- [32] International Quality and Productivity Center corporate website. <http://www.iqpc.com/cgi-bin/templates/document.html?topic=483&event=5992&document=47445>
- [33] "Parco Merged Media has received a contract from the Washington Hospital Center for installation of its wireless tracking system throughout the emergency department," Parco press release, August 11, 2004. <http://www.parcomergedmedia.com/scripts/pr.asp?Action=Show&ID=38>
- [34] Parco Merged Media corporate website. <http://www.parcowireless.com/>
- [35] Collins, J., "Hospital gets ultra-wideband RFID," *RFID Journal*, August 19, 2004. <http://www.rfidjournal.com/article/view/1088>
- [36] *RFID Journal*. <http://www.rfidjournal.com/>
- [37] *IDTechEx*. <http://www.idtechex.com/>
- [38] Patterson, D., Liao, L., Fox, D., and Kautz, H., "Inferring high-level behavior from low-level sensors," in *UbiComp 2003*, Springer-Verlag, LNCS 2864, pp. 73-89, 2003.
- [39] Patterson, D., Etzioni, O., Fox, D., Kautz, H., "The Activity Compass." In *Proceedings of UBICOG '02: First International Workshop on Ubiquitous Computing for Cognitive Aids*, 2002.
- [40] Doucet, A., de Freitas, N., Gordon, N., eds., *Sequential Monte Carlo in Practice*, Springer-Verlag, New York, 2001.
- [41] Liao, L., Fox, D., Hightower, J., Kautz, H., Schulz, D., "Voronoi tracking: Location estimation using sparse and noisy sensor data," In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2003.
- [42] Bar-Shalom, Y., Li, X., Kirubarajan, T., *Estimation with Applications to Tracking and Navigation*, John Wiley, 2001.
- [43] Fishkin, K., Bing, J., Philipose, M., and Roy, S., "I sense a disturbance in the force: Unobtrusive detection of interactions with RFID-tagged objects," in *UbiComp 2004*, Springer-Verlag, LNCS 3205, pp. 268-282, 2004.
- [44] Munguia, E., *Activity Recognition in the Home Setting Using Simple and Ubiquitous Sensors*, MIT M.S. Thesis, 2003.
- [45] Munguia, E., Intille, S., and Larson, K., "Activity recognition in the home using simple and ubiquitous sensors," in *Pervasive 2004*.

- [46] Morris, M., Lundell, J., Dishman, E., and Needham, B., "New perspectives on ubiquitous computing from ethnographic study of elders with cognitive decline," in *UbiComp 2003*, Springer-Verlag, LNCS 2864, pp. 227-242, 2003.
- [47] Salvador, T., and Anderson, K., "Practical considerations of context for context based systems: An example from an ethnographic case study of a man diagnosed with early onset Alzheimer's disease," in *UbiComp 2003*, Springer-Verlag, LNCS 2864, pp. 243-255, 2003.
- [48] Degani, A., "On the types of modes in human-machine interactions," in *Proceedings of The Ninth international Symposium on Aviation Psychology*, Columbus, Ohio, 28 April -1 May 1997.
- [49] Rudisill, M., "Line pilots' attitudes about and experience with flight deck automation: Results of an international survey and proposed guidelines," in *Proceedings of the Eighth International Symposium on Aviation Psychology*, Columbus, Ohio, 1995.
- [50] Sarter, N., and Woods, D., "Team play with a powerful and independent agent: Operational experiences and automation surprises on the Airbus A-320," *Human Factors* 39, no. 4 (December 1997), 558.
- [51] Christensen, H., and Bardram, J., "Supporting human activities—Exploring activity-centered computing," in *UbiComp 2002*, Springer-Verlag, LNCS 2498, pp. 107-116, 2002.

**Exhibit 1-B**

**UMMC  
Indoor Positioning System  
Evaluation Facility**

**5 July 2005**

**J. Naslund, UMMC  
B. Bame, UMMS  
D. Clarke, Fremont Associates, LLC**

## **1. Introduction**

The use of systems for passively tracking the location of objects and people at the room or building level is expanding rapidly in the healthcare domain. These *Indoor Positioning Systems* (IPS) are installed to enhance key health care performance indicators like workflow, utilization of resources, and patient safety. Effective deployment of an IPS, and consequent improvement in key indicators, is dependent on the technical characteristics of the deployed system, and the ability of the system to operate in its intended environment.

As part of its Operating Room of the future project, the University of Maryland Medical Center has established an Indoor Positioning System Evaluation Facility. It consists of a testing methodology, supplies and instrumentation, and dedicated operating rooms for the evaluation of the full range of indoor positioning system technologies. The goal of the facility is to experiment with commercial, off-the-shelf systems to evaluate their technical characteristics and fitness for use in the perioperative environment. The results of system evaluations will be made public.

The remainder of this document describes the facilities, methodology and detailed test procedures to be employed. Section 2 will describe the resources allocated for the evaluation facility, including physical space and equipment. Section 3 presents the high-level test plan, outlining the standard configuration and test criteria to be employed. Section 4 will describe our plans for disseminating results of evaluation activities. Section 5 closes this document with summary remarks. Detailed test scripts for evaluation of systems are included as an Appendix.

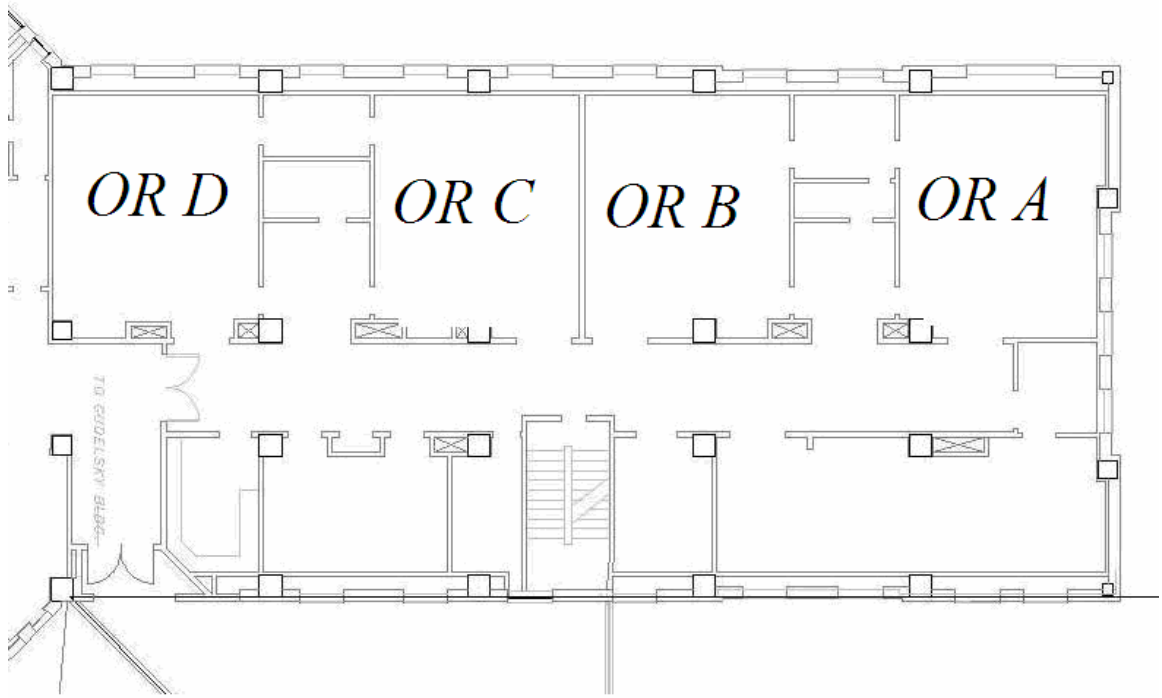
## **2. Facilities**

The IPS Evaluation Facility is made up of staff with expertise in IPS technologies, dedicated physical space for conducting evaluation activities, and an assemblage of infrastructure (computing hardware, networking hardware, *etc.*), healthcare equipment and tools and supplies.

### **a. Physical Space**

The IPS Evaluation Facility is housed in the UMMC Simulation Center, on the seventh floor of the Swirnow building. The Simulation Center is being constructed in a former UMMC operating room suite. These were active medical center operating rooms until operating rooms were placed into service in 2003. The floor plan for the simulation center is presented in Figure 5.

As shown in Figure 5, the Simulation Center consists of a long corridor separating the four operating rooms from general purpose wet/dry lab and office space. The IPS Evaluation Facility will make use of OR D and OR C and the service areas between them, as well as OR A. OR's D and C will provide space for deployment of systems and their evaluation under static conditions. OR A, which houses a Vicon motion capture system, will be used for dynamic tests involving the detection of tags in motion.



**Figure 5: UMMC Simulation Center**

## **b. Equipment**

Evaluation Center equipment is grouped into three major categories: Infrastructure items, healthcare equipment and tools and supplies.

### **1. Infrastructure**

- a. Laptop with 802.11 b/g wireless NIC
- b. Desktop PC with ample disk space, Gb NIC, CD writer
- c. Furniture (desk, chair)
- d. 16 Port 10/100 switch
- e. As-built engineering drawings of the test environment
- f. Vicon motion capture system
- g. 802.11 a/b/g access point hardware (Quantity: as needed)

Infrastructure items consist of basic furniture, computing equipment and networking gear to support the activities of the evaluation center (a-d), as-built engineering drawings of the test environment for calibrating IPS location software (e), a Vicon motion capture system for establishing ground truth location data for tags while in motion (f), and WiFi networking hardware to support 802.11-based IPS system evaluation (g).

The Vicon motion capture system is a high precision system for determining the location of fiducial markers in three dimensions. A system of 12 independent cameras with infrared light sources and detectors are used to capture the room-level scene at a sampling rate of 100 samples per second. The captured data can be used to determine position in three dimensions to a very high degree of accuracy. Our tests in the OR A laboratory

demonstrated that accuracy was within 0.2% of manually measured values. This equates to less than one centimeter of measurement error across a 4 meter wide room.

## 2. Healthcare equipment

- a. Suture cart
- b. Scanner
- c. Monitor cart
- d. Packaged individual instruments (i.e. scalpels, scopes, clamps)
- e. Unpackaged individual instruments (i.e. scalpels, scopes, clamps)
- f. Instrument casing with instruments included
- g. Instrument table
- h. Gurney

Healthcare equipment encompasses items typically found in the perioperative setting. These will be used to evaluate the ease of mounting tags, as well as determining the impact of mounting on the accuracy of reported position data.

## 3. Tools and Supplies

- a. Power strips
- b. AC Extension cords
- c. Batteries
- d. Cat5/RJ45 crimping tool
- e. LACK Shelving Unit (Ikea) (Quantity: 5)
  - i. 16"x10" right angle bracket (Quantity: 10)
  - ii. Wood screws (Quantity: 30)
  - iii. Bolts (Quantity: 10)
- f. Measuring tape
- g. Masking tape
- h. Stopwatch
- i. Wire strippers
- j. Blank CD's
- k. Postal scale
- l. Velcro straps
- m. Wire ties

Tools and supplies are basic items to facilitate set-up, evaluation and tear-down activities in the evaluation facility.

## 3. Test Plan

Systems will be evaluated according to pre-established criteria. The key elements of these criteria are a standard configuration framework and evaluation criteria.

### a. Standard Configuration

*OR's D and C for static tests, OR A for dynamic tests*—All evaluation activities will be carried out in the established test areas. Static tests of location detection accuracy will be performed in OR's D and C and the

service areas between them. Dynamic tests will be carried out in OR A, where the Vicon motion tracking system is located.

*Vendor specified reader configuration*—We will solicit and employ a vendor-specified reader configuration for the evaluation areas. Vendors will be provided with the as-built drawings and test criteria and asked for guidance on the appropriate quantity and placement of readers.

*15 tags*—Tests will employ a maximum of 15 tags. This limit applies to a system's ability to report co-location, as well as a system's ability to tolerate multiple tags in a single reader's read zone.

*802.3 10/100 network*—Where 802.3 network interconnection between readers and server machines is required we will employ a dedicated 10/100 network switch.

*Integrated server configuration (per vendor approval)*—Vendor supplied software will be run on a single dedicated server computer, regardless of how many applications, tasks, *etc.*, are to be run. The soundness of this configuration will be verified with vendors before testing.

## **b. Evaluation Criteria**

The general evaluation criteria focus on three major areas: physical characteristics of the systems (readers and tags), the accuracy of position data reported, and system characteristics.

1. Physical characteristics
  - a. Size, weight, *etc.*
  - b. Maintenance
  - c. Tolerance of cleaning, sterilization, fluids, *etc.*

Physical characteristics of each system will be assessed through inspection, consultation with the vendor, and experimentation. The size, weight, *etc.*, of tags and readers will be assessed to determine the burden their carrying or mounting will entail.

Maintenance needs of tags and readers—batteries, cleaning, calibration, *etc.*—will be ascertained through inspection of items and their documentation, and through consultation with the vendor.

The tolerance of tags for solutions and processes that may be encountered in the perioperative environment will be assessed first by consultation with vendors. Where claims of tolerance are made experiments will be carried out to verify these claims. The methods of treatment and solutions to which tags will be exposed will be determined based on current UMMC practices as well as AAMI sterilization standards.

2. Accuracy
  - a. Static
  - b. Dynamic
  - c. Co-location



Tests of the accuracy of reported location data will be carried out under conditions both static (*i.e.*, tags in fixed locations) and dynamic (*i.e.*, tags in motion during measurements). In addition, we will carry out static and dynamic tests to evaluate the ability of systems to differentiate co-located items.

Static tests will be carried out by placing tags at a wide range of pre-defined (and pre-measured) locations in the OR D/C areas and recording their reported location. Test scenarios will involve various confounding factors such as attachment to various materials, materials blocking the line of sight between the tag and one or more readers, *etc.*

Dynamic tests will be carried out by moving tags through the Vicon system's imaging area at rates of speed within predetermined bounds. As in the static case, various confounding factors will be added to assess their impact on reported accuracy.

Tests of the systems' abilities to distinguish co-located tags will be carried out by placing sensors as close to each other as possible. The system's reported location data will be recorded as the sensors are moved away from one another in pre-defined increments.

### 3. System characteristics

- a. Tolerance for multiple tags in single read zone
- b. Integration effort
- c. RF characteristics
- d. Interoperability
- e. Privacy/Security
- f. Data communication load
- g. Sensitivity to battery strength
- h. Sampling rate
- i. Tag storage/transmission features
- j. Bleed through

A wide range of system-level tests will be conducted to assess each system's overall characteristics with respect to integration, customization, usability, *etc.*

To evaluate a system's tolerance for multiple tags in a single read zone varying numbers of tags will be placed within a read zone. The resulting location data will be evaluated to determine which tags were sensed and how accurate the reported position information is.

The effort involved in using the position data reported by an IPS to drive another application will be assessed through consultation and inspection. Vendor representatives will be consulted for information on the suggested approach to integration. Any documentation or API resources provided by the vendor will be inspected and evaluated.

The RF emissions of the system will be ascertained from documentation and consultation with the vendor. These emissions will be compared with the known uses of RF spectrum by existing systems in the medical center.

The ability of a vendor's system to operate in the same physical space as other vendors' equipment will be evaluated by concurrent static testing. The quality of reported data under these conditions will be assessed.

The privacy and security of a system will be evaluated by consultation with the vendor and inspection of documentation. Results will be reported for a number of key factors impacting privacy and security.

The data communication load generated by a vendor's system will be assessed through consultation with the vendor and analytic modeling.

Sensitivity to battery strength will be assessed through static tests. Batteries will be drained to pre-set levels and re-used to determine any impact on the accuracy of reported results.

Sampling rate will be ascertained through consultation with the vendor and through inspection of documentation. Sampling rate will be objectively measured through testing. The deviation from the expected sampling rate will be assessed.

The storage and transmission capability of the tags will be assessed.

Detailed test scripts for these criteria are included in the appendix.

#### **4. Dissemination**

Results will be shared with TATRC as part of the normal reporting process for UMMC's Operating Room of the Future project. Results will be publicly disseminated through the UMMC ORF website at <TBD>.

#### **5. Summary**

We have described facilities and test criteria for the establishment of an Indoor Positioning System Evaluation Facility as part of the University of Maryland Medical Center's Operating Room of the Future project. The goal of this effort is to provide independent evaluation of COTS indoor positioning systems. These evaluations will be carried out in a way that emphasizes the unique constraints of the perioperative environment. Though the core of the testing methodology is focused on repeating identical tests for each vendor's system, we also intend to assess and report the unique features of each vendor's offering.

Resources have been allocated for this project and the first IPS system evaluations will be completed during the summer of 2005. We expect evaluation activities to continue throughout the UMMC ORF contract performance period.

## **Appendix: Test Scripts**

The pages that follow present detailed scripts for repeatable tests to be carried out on each system evaluated. Where possible, results will be determined through well-defined objective experimentation. When subjective evaluation is necessary, the evaluation criteria will be described in detail in advance.

There are a total of <?> test scripts covering three major areas: physical characteristics of the systems (readers and tags), the accuracy of position data reported, and system characteristics.

## **Test Setup**

This section describes the steps necessary to setup the vendor's system for testing. This step will be performed once for each system. Tests that require a fully setup system will have "Fully-functional system" listed in the resources needed section and will have a step called "setup vendors system".

**Test Number:** N/A

**Test Name:** System setup

**Author:** JMN

**Last Updated:** 21 July 2005

**Estimated time for completion:** 8 hours

### **Resources needed:**

1. Laptop with 802.11 b/g wireless NIC
2. Desktop PC with ample disk space, Gb NIC, CD writer
3. 16 Port Gb switch
4. As-built engineering drawings of the test environment
5. 802.11 a/b/g access point hardware
6. Power strips
7. AC Extension cords
8. Cat5/RJ45 crimping tool
9. Measuring tape
10. Masking tape
11. Wire ties
12. System documentation
13. Vendor consultant (optional)

### **Materials consumed:**

1. Masking tape
2. Wire ties
3. Cat5 cable

### **Detailed steps:**

#### **1. Preparation:**

- a. Gather required resources

#### **2. Execution:**

- a. Place readers in the optimal locations at ceiling height

- b. Connect readers to LAN with Cat5 cable
  - c. Connect readers to power strips
  - d. Connect power strips to outlets
  - e. Connect laptop to LAN with Cat5 cable
  - f. Connect desktop to LAN with Cat5 cable
  - g. Consolidate cables with wire ties
  - h. Calibrate system based on documentation and vendor consultant advice
  - i. Configure software based on documentation and vendor consultant
2. **Expected result:**
- a. Fully functional IPS

## **Test Breakdown**

This section describes the steps necessary to breakdown a vendor's system after testing has been completed. This step will be performed once for each system. This step will be performed once all tests for a given system have been completed.

**Test Number:** N/A

**Test Name:** System breakdown

**Author:** JMN

**Last Updated:** 21 July 2005

**Estimated time for completion:** 4 hours

### **Resources needed:**

- 1. Original packaging from vendor's system

### **Materials consumed:**

- 1. Masking tape

### **Detailed steps:**

- 1. **Preparation:**
  - a. Gather required resources
- 2. **Execution:**
  - a. Place all components of vendor's system in original packaging
  - b. Return packaging to storage
- 3. **Expected result:**
  - a. A packaged system in an inventoried location

## **Note**

The estimated time for completion is based on a "generic" system. Some systems may have features or limitations that cause certain tests to require considerably more or less time to complete. These issues will be addressed "on the fly". Minor adjustments may be made to test scripts for certain systems.

**Test Number:** P1.1.1

**Test Name:** Tag weight

**Author:** DEC

**Last Updated:** 7 July 2005

**Major Criterion:** Physical characteristics

**Minor Criterion:** Size, weight, *etc.*

**Estimated time for completion:** 15 minutes

**Scenario:** Weigh a tag in its fully functional configuration.

**Resources needed:**

1. Tag
2. Postal scale

**Materials consumed:**

None

**Detailed steps:**

**1. Preparation:**

- a. Gather required resources
- b. Verify that all components (especially battery) are present in tag

**2. Execution:**

- a. Place tag on scale.
- b. Record weight (including units) in log

**3. Expected result:**

- a. A quantity “reasonable” for the given tag

**4. Evaluation:**

- a. None; this is an objective measure.

**5. Reporting:**

- a. Add to table of per vendor, per tag weights
- b. Note any special features/constraints of the tag

**6. Clean-up:**

- a. Return resources to storage

**Test Number:** P1.2.1

**Test Name:** Reader weight

**Author:** JMN

**Last Updated:** 10 July 2005

**Major Criterion:** Physical characteristics

**Minor Criterion:** Size, weight, *etc.*

**Estimated time for completion:** 15 minutes

**Scenario:** Weigh a reader in its fully functional configuration.

**Resources needed:**

1. Reader
2. Postal scale

**Materials consumed:**

None

**Detailed steps:**

**1. Preparation:**

- a. Gather required resources
- b. Verify that all components are present in reader

**2. Execution:**

- a. Place reader on scale.
- b. Record weight (including units) in log

**3. Expected result:**

- a. A quantity “reasonable” for the given reader

**4. Evaluation:**

- a. None; this is an objective measure

**5. Reporting:**

- a. Add to table of per vendor, per reader weights
- b. Note any special features/constraints of the reader

**6. Clean-up:**

- a. Return resources to storage

**Test Number:** P1.3.1

**Test Name:** Tag size

**Author:** JMN

**Last Updated:** 10 July 2005

**Major Criterion:** Physical characteristics

**Minor Criterion:** Size, weight, *etc.*

**Estimated time for completion:** 15 minutes

**Scenario:** Measure a tag's dimensions (length, width height)

**Resources needed:**

1. Tag
2. Measuring tape

**Materials consumed:**

None

**Detailed steps:**

**1. Preparation:**

- a. Gather required resources
- b. Verify that all components are present on tag

**2. Execution:**

- a. Measure tag's length
- b. Record length in log
- c. Measure tag's width
- d. Record width in log
- e. Measure tag's height
- f. Record width in log

**3. Expected result:**

- a. 3 quantities "reasonable" for the given tag.

**4. Evaluation:**

- a. None; this is an objective measure.

**5. Reporting:**

- a. Add to table of per vendor, per tag dimensions.
- b. Note any special features/constraints of the tag.

**6. Clean-up:**

- a. Return resources to storage

**Test Number:** P1.4.1

**Test Name:** Reader size

**Author:** JMN

**Last Updated:** 10 July 2005

**Major Criterion:** Physical characteristics

**Minor Criterion:** Size, weight, *etc.*

**Estimated time for completion:** 15 minutes

**Scenario:** Measure a reader's dimensions (length, width height)

**Resources needed:**

3. Reader
4. Measuring tape

**Materials consumed:**

None

**Detailed steps:**

**1. Preparation:**

- a. Gather required resources
- b. Verify that all components are present in reader

**2. Execution:**

- a. Measure reader's length
- b. Record length in log
- c. Measure reader's width
- d. Record width in log
- e. Measure reader's height
- f. Record width in log

**3. Expected result:**

- a. 3 quantities "reasonable" for the given reader

**4. Evaluation:**

None; this is an objective measure

**5. Reporting:**

- a. Add to table of per vendor, per reader dimensions
- b. Note any special features/constraints of the reader

**6. Clean-up:**

- a. Return resources to storage



**Test Number:** P1.5.1

**Test Name:** Reader layout

**Author:** JMN

**Last Updated:** 20 July 2005

**Major Criterion:** Physical characteristics

**Minor Criterion:** Size, weight, *etc.*

**Estimated time for completion:** 15 minutes

**Scenario:** Examine a reader and note areas that, if obstructed, may reduce accuracy. Run test to determine effect on accuracy.

**Resources needed:**

1. Reader

**Materials consumed:**

None

**Detailed steps:**

**1. Preparation:**

- a. Gather required resources
- b. Verify that all components are present on reader

**2. Execution:**

- a. Note areas that, if obstructed, may reduce accuracy in log
- b. Obstruct the noted areas with cardboard. Affix cardboard with masking tape
- c. Perform A\*.\*.\* tests

**3. Expected result:**

- a. Areas that, if obstructed, may reduce accuracy
- b. Effect of obstruction on accuracy

**4. Evaluation:**

- a. None; this is an objective measure

**5. Reporting:**

- a. Add to table of per vendor areas that, if obstructed, may reduce accuracy
- b. Add the results of the A\*.\*.\* tests to the log

**6. Clean-up:**

- a. Return resources to storage

**Test Number:** P1.6.1

**Test Name:** Reader mounting capability

**Author:** JMN

**Last Updated:** 20 July 2005

**Major Criterion:** Physical characteristics

**Minor Criterion:** Size, weight, *etc.*

**Estimated time for completion:** 15 minutes

**Scenario:** Evaluate the reader's mounting capabilities through examination of documentation and through examination of the reader.

**Resources needed:**

1. Reader
2. Reader specifications/manual
3. Any additional vendor supplied equipment for mounting a reader

**Materials consumed:**

None

**Detailed steps:**

**1. Preparation:**

- a. Gather required resources
- b. Verify that all components are present on reader

**2. Execution:**

- a. Examine reader; note mounting options
- b. Examine reader documentation; note mounting options

**3. Expected result:**

- a. Mounting options for given reader

**4. Evaluation:**

- a. None; this is an objective measure

**5. Reporting:**

- a. Add to table of per vendor, per reader mounting options
- b. Note any special equipment required for mounting

**6. Clean-up:**

- a. Return resources to storage

**Test Number:** P1.7.1

**Test Name:** Tag mounting capability

**Author:** JMN

**Last Updated:** 20 July 2005

**Major Criterion:** Physical characteristics

**Minor Criterion:** Size, weight, *etc.*

**Estimated time for completion:** 15 minutes

**Scenario:** Evaluate the tag's mounting capabilities through examination of documentation and through examination of the tag

**Resources needed:**

1. Tag
2. Tag specifications/manual
3. Any additional vendor supplied equipment for mounting a tag

**Materials consumed:**

None

**Detailed steps:**

**1. Preparation:**

- a. Gather required resources
- b. Verify that all components are present on tag

**2. Execution:**

- a. Examine tag; note mounting options
- b. Examine tag documentation; note mounting options

**3. Expected result:**

- a. Mounting options for given tag

**4. Evaluation:**

- a. None; this is an objective measure

**5. Reporting:**

- a. Add to table of per vendor, per tag mounting options
- b. Note any special equipment required for mounting

**6. Clean-up:**

- a. Return resources to storage

**Test Number:** P2.1.1

**Test Name:** Tag maintenance

**Author:** JMN

**Last Updated:** 20 July 2005

**Major Criterion:** Physical characteristics

**Minor Criterion:** Maintenance

**Estimated time for completion:** 15 minutes

**Scenario:** Determine the tag's maintenance requirements.

**Resources needed:**

1. Tag specifications/manual
2. Vendor representative

**Materials consumed:**

None

**Detailed steps:**

**1. Preparation:**

- a. Gather required resources

**2. Execution:**

- a. Examine documentation to determine tag's maintenance requirements.
- b. Consult system representative to determine tag's maintenance requirements.
- c. Record maintenance requirements in log

**3. Expected result:**

- a. Maintenance requirements for the given tag. A maintenance task is defined as: A task that must performed (by a human) at a regular interval to ensure optimal function. Such tasks include but are not limited to
  - a. Physical cleaning
  - b. Software update
  - c. Firmware update
  - d. Hardware update
  - e. Software recalibration
  - f. Hardware recalibration
  - g. Software backup
  - h. Hardware replacement
  - i. System evaluation

**4. Evaluation:**

- a. None; this is an objective measure

**5. Reporting:**

- a. Add to table of per vendor, per tag maintenance requirements. This data will consist of:
  - a. A table with the following columns (each row will be a maintenance task)
    - i. Time maintenance must be performed in days
    - ii. Maintenance to be performed

- iii. Time required to complete maintenance
- iv. Cost to complete maintenance in dollars
- v. Personnel required to complete maintenance
- vi. Effect of failure to perform task

**6. Clean-up:**

- a. Return resources to storage

**Test Number:** P2.2.1

**Test Name:** Tag battery life

**Author:** JMN

**Last Updated:** 10 July 2005

**Major Criterion:** Physical characteristics

**Minor Criterion:** Maintenance

**Estimated time for completion:** 15 minutes

**Scenario:** Determine the tag's battery life.

**Resources needed:**

1. Tags specifications/manual
2. Vendor representative

**Materials consumed:**

None

**Detailed steps:**

1. **Preparation:**
  - a. Gather required resources
2. **Execution:**
  - a. Determine tag's battery life
  - b. Determine time required to recharge tag
  - c. Record battery life in log.
3. **Expected result:**
  - a. Battery life in hours for the given tag
  - b. Time required to recharge tag in hours
4. **Evaluation:**
  - a. None; this is an objective measure.
5. **Reporting:**
  - a. Add to table of per vendor, per tag battery life
  - b. Add to table of per vendor, per tag recharge time
6. **Clean-up:**
  - a. Return resources to storage

**Test Number:** P2.3.1

**Test Name:** Tag alert function

**Author:** JMN

**Last Updated:** 10 July 2005

**Major Criterion:** Physical characteristics

**Minor Criterion:** Maintenance

**Estimated time for completion:** 15 minutes

**Scenario:** Determine the tag's ability to "alert" a user/technician of maintenance needs

**Resources needed:**

1. Tags specifications/manual
2. Vendor representative

**Materials consumed:**

None

**Detailed steps:**

1. **Preparation:**
  - a. Gather required resources
2. **Execution:**
  - a. Determine tag's ability to alert
  - b. Record tag's ability to alert in log.
3. **Expected result:**
  - a. Tag's ability to alert
4. **Evaluation:**
  - a. None; this is an objective measure
5. **Reporting:**
  - a. Add to table of per vendor, per tag alert ability
6. **Clean-up:**
  - a. Return resources to storage

**Test Number:** P2.5.1

**Test Name:** Reader maintenance

**Author:** JMN

**Last Updated:** 20 July 2005

**Major Criterion:** Physical characteristics

**Minor Criterion:** Maintenance

**Estimated time for completion:** 15 minutes

**Scenario:** Determine the reader's maintenance requirements

**Resources needed:**

1. Reader specifications/manual
2. Vendor representative

**Materials consumed:**

None

**Detailed steps:**

**1. Preparation:**

- a. Gather required resources

**2. Execution:**

- a. Examine documentation to determine reader's maintenance requirements.
- b. Consult system representative to determine reader's maintenance requirements
- c. Record maintenance requirements in log

**3. Expected result:**

- a. Maintenance requirements for the given reader. A maintenance task is defined as: A task that must performed (by a human) at a regular interval to ensure optimal function. Such tasks include but are not limited to:
  - i. Physical cleaning
  - ii. Software update
  - iii. Firmware update
  - iv. Hardware update
  - v. Software recalibration
  - vi. Hardware recalibration
  - vii. Software backup
  - viii. Hardware replacement
  - ix. System evaluation

**4. Evaluation:**

- a. None; this is an objective measure

**5. Reporting:**

- a. Add to table of per vendor, per reader maintenance requirements. This data will consist of:
  - i. A table with the following columns (each row will be a maintenance task)
    1. Time maintenance must be performed in days
    2. Maintenance to be performed



3. Time required to complete maintenance
4. Cost to complete maintenance in dollars
5. Personnel required to complete maintenance
6. Effect of failure to perform task

**6. Clean-up:**

- a. Return resources to storage

**Test Number:** P2.5.1

**Test Name:** Maintenance cost

**Author:** JMN

**Last Updated:** 10 July 2005

**Major Criterion:** Physical characteristics

**Minor Criterion:** Maintenance

**Estimated time for completion:** 15 minutes

**Scenario:** Determine the maintenance costs of the system

**Resources needed:**

1. System specifications/manual
2. Results of P2.1.1, P2.4.1

**Materials consumed:**

None

**Detailed steps:**

**1. Preparation:**

- a. Gather required resources

**2. Execution:**

- a. Determine maintenance costs of the system
- b. Record maintenance requirements in log

**3. Expected result:**

- a. Maintenance costs “reasonable” for the given reader

**4. Evaluation:**

- a. None; this is an objective measure

**5. Reporting:**

Add to table of per vendor maintenance costs

**6. Clean-up:**

- a. Return resources to storage

**Test Number:** P2.6.1

**Test Name:** Calibration evaluation

**Author:** JMN

**Last Updated:** 10 July 2005

**Major Criterion:** Physical characteristics

**Minor Criterion:** Maintenance

**Estimated time for completion:** 30 minutes

**Scenario:** Determine the calibration requirements of the system.

**Resources needed:**

1. All equipment supplied by the vendor
2. Measuring tape
3. Duct tape
4. Laser leveler/rangefinder
5. Misc. equipment needed for calibration
6. Vendor representative

**Materials consumed:**

None

**Detailed steps:**

**1. Preparation:**

- a. Gather required resources

**2. Execution:**

- a. Calibrate the system based on documentation and vendor representative's guidance.
- b. Add each step performed to log
- c. Add time to complete each step to log

**3. Expected result:**

- a. Calibrated system
- b. Log of calibration

**4. Evaluation:**

None; this is an objective measure

**5. Reporting:**

- a. Add to per vendor data for calibration evaluation. This data will consist of the following:

- i. Table of calibration steps

Each row in the table will represent a step. The table has the following columns:

1. Step name
2. Step description
3. Time required
4. Personnel required

**6. Clean-up:**

- a. Return resources to storage.

**Test Number:** P3.1.1

**Test Name:** Sterilization (Heat)

**Author:** JMN

**Last Updated:** 10 July 2005

**Major Criterion:** Physical characteristics

**Minor Criterion:** Tolerance of cleaning, sterilization, fluids, *etc.*

**Estimated time for completion:** 4 hours

**Scenario:** Determine the tags resistance to sterilization.

**Resources needed:**

1. Tag
2. Access to Autoclave
3. Sterilization staff member

**Materials consumed:**

Tag (if destroyed)

**Detailed steps:**

**1. Preparation:**

- a. Gather required resources
- b. Confirm tag's tolerance with vendor

**2. Execution (if vendor did not confirm tolerance the test will not be executed and the result will be an automatic "No"):**

- a. Place the tag in the Autoclave
- b. Run the Autoclave at the normal settings. (115 Celsius for 15 minutes)
- c. Record any physical abnormalities caused by sterilization.
- d. Bring the tag back to the Sim Center
- e. Test the tags function

**3. Expected result:**

- a. Yes/No (yes if the tag still functions after sterilization, no if it no longer functions after sterilization)
- b. Physical deformations caused by sterilization

**4. Evaluation:**

- a. Place the tag into one of the following categories based on the observed physical discolorations
  - i. No change
  - ii. Slight discoloration
  - iii. Moderate discoloration
  - iv. Severe discoloration
- b. Place tag into one of the following categories based on the observed physical deformations
  - i. No change
  - ii. Bubbling
  - iii. Warped/melted; retaining original shape
  - iv. Warped/melted; lump or blob

**5. Reporting:**

- a.** Add to per vendor data for sterilization (heat) test. This data will consist of the following:
  - i.** Yes/No (yes if the tag still functions after sterilization, no if it no longer functions after sterilization)
  - ii.** Placement in discoloration scale
  - iii.** Placement in deformation scale

**6. Clean-up:**

Return resources to storage

**Test Number:** P3.2.1

**Test Name:** Sterilization (Chemical)

**Author:** JMN

**Last Updated:** 10 July 2005

**Major Criterion:** Physical characteristics

**Minor Criterion:** Tolerance of cleaning, sterilization, fluids, *etc.*

**Estimated time for completion:** 4 hours

**Scenario:** Determine the tags resistance to sterilization.

**Resources needed:**

1. Tag
2. Access to Peracetic acid sterilization device
3. Sterilization staff member

**Materials consumed:**

Tag (if destroyed)

**Detailed steps:**

**1. Preparation:**

- a. Gather required resources

**2. Execution:**

- a. Place the tag in the sterilization device
- b. Run the sterilization process at the normal settings
- c. Record any physical abnormalities caused by sterilization
- d. Bring the tag back to the Sim Center
- e. Test the tags function

**3. Expected result:**

- a. Yes/No (yes if the tag still functions after sterilization, no if it no longer functions after sterilization)
- b. Physical abnormalities caused by sterilization

**4. Evaluation:**

- a. Place the tag into one of the following categories based on the observed physical discolorations
  - i. No change
  - ii. Slight discoloration
  - iii. Moderate discoloration
  - iv. Severe discoloration
- b. Place tag into one of the following categories based on the observed physical corrosion
  - i. 100% of tag remaining
  - i. 75% of tag remaining
  - ii. 50% of tag remaining
  - iii. 0% of tag remaining

**5. Reporting:**

- a. Add to per vendor data for sterilization (chemical) test. This data will consist of the following

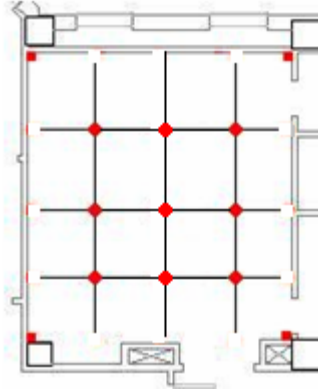
- i. Yes/No (yes if the tag still functions after sterilization, no if it no longer functions after sterilization)
- ii. Placement in discoloration scale
- iii. Placement in corrosion scale

**6. Clean-up:**

Return resources to storage

**Test Number:** A1.1.1  
**Test Name:** Static location  
**Author:** JMN  
**Last Updated:** 10 July 2005

**Major Criterion:** Accuracy  
**Minor Criterion:** Static  
**Estimated time for completion:** 5 hours



**Scenario:** The OR will be divided into a 3x3 grid as shown in the diagram. A tag will be placed at each location (red dot) and its reported location recorded for 30 seconds at the system's highest refresh rate.

**Resources needed:**

1. Tag
2. Measuring tape
3. Readers (Quantity: vendor-specified)
4. Laptop with 802.11 b/g wireless NIC
5. Desktop PC with ample disk space, Gb NIC, CD writer
6. As-built engineering drawings of the test environment
7. Fully functional system
8. Misc. equipment required by vendor's system

**Materials consumed:**

1. Masking tape

**Detailed steps:**

**1. Preparation:**

- a. Gather required resources
- b. Setup vendor's system
- c. Measure 3x3 grid; mark with masking tape
- d. Verify that all components of the system are functioning properly

**2. Execution:**

- a. Place a tag at a location. The tag will be placed on a table at waist height.
- b. Record the tags reported location for 15 seconds at the system's highest refresh rate
- c. Repeat at each location

**3. Expected result:**

- a. Reported location data at each location. Number of samples:  $15 \times \text{system's refresh rate/sec}$

**4. Evaluation:**

None; this is an objective measure

**5. Reporting:**

- a. Add to per vendor data for static location results. This data will consist of the following:
  - i. Table of location data  
Each row will represent one sample (total of 15). The table has the following columns:



1. Timestamp in ms
2. Reported X
3. Reported Y
4. True X
5. True Y
6. Percentage error in X
7. Percentage error in Y

**6. Clean-up:**

- a. Return resources to storage

**Test Number:** A1.2.1

**Test Name:** Line of sight (Static)

**Author:** JMN

**Last Updated:** 10 July 2005

**Major Criterion:** Accuracy

**Minor Criterion:** Static

**Estimated time for completion:** 4 hours

**Scenario:** Place a tag in the center of OR1 and record its reported location while varying the number of readers that have line of sight (LOS) to the tag.

**Resources needed:**

1. Tag
2. Measuring tape
3. Readers (Quantity: vendor-specified)
4. Large cabinets (Quantity: equal to number of readers)
5. Laptop with 802.11 b/g wireless NIC
6. Desktop PC with ample disk space, Gb NIC, CD writer
7. As-built engineering drawings of the test environment
8. Fully functional system
9. Misc. equipment required by vendor's system

**Materials consumed:**

1. Masking tape

**Detailed steps:**

**1. Preparation:**

- a. Gather required resources
- b. Setup vendor's system
- c. Measure center of room; mark with duct tape. Tag will be placed on a table at waist height
- d. Place cabinet(s) against the walls so they do not block LOS
- e. Verify that all components of the system are functioning properly

**2. Execution:**

- a. Place tag in center of room. (Wall length / 2 away from 2 adjacent walls)
- b. Record the tag's reported location for 15 seconds at the system's highest refresh rate
- c. Block LOS on one of the readers and repeat
- d. Continue repeating until the test is run with all readers blocked.

**3. Expected result:**

- a. Reported location data for each number of readers blocked. (0 to reader density) Number of samples: 15\*system's refresh rate/sec

**4. Evaluation:**

None; this is an objective measure

**5. Reporting:**

- a. Add to the per vendor data for line of sight results. This data will consist of the following:
  - i. A table of results for each number of readers blocked. (0 to reader density)

Each row will represent one sample (total of 15). The table has the following columns:

1. Timestamp in ms
2. Reported X
3. Reported Y
4. True X
5. True Y
6. Percentage error in X
7. Percentage error in Y

**6. Clean-up:**

- a. Return resources to storage

**Test Number:** A1.3.1

**Test Name:** Contrived OR (Static)

**Author:** JMN

**Last Updated:** 10 July 2005

**Major Criterion:** Accuracy

**Minor Criterion:** Static

**Estimated time for completion:** 5 hours

**Scenario:** Construct a mock OR that replicates that obstructions and materials that will be present in a real OR.

**Resources needed:**

1. Tag
2. Measuring tape
3. Readers (Quantity: vendor-specified)
4. Large cabinets
5. Laptop with 802.11 b/g wireless NIC
6. Desktop PC with ample disk space, Gb NIC, CD writer
7. As-built engineering drawings of the test environment
8. Fully functional system
9. Misc. equipment required system.
10. Suture cart
11. Scanner
12. Monitor cart
13. Instrument table
14. Gurney
15. Human
16. Human

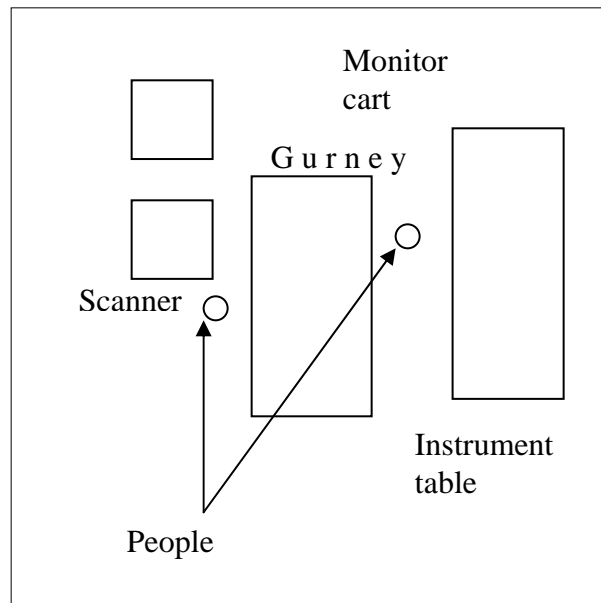
**Materials consumed:**

1. Masking tape

**Detailed steps:**

**1. Preparation:**

- a. Gather requires resources
- b. Setup vendor's system
- c. Tag the following objects and place the following objects at known locations near the center of the room (see diagram). Their arrangement should resemble a true OR:
  - i. Gurney (tag laid on top)
  - ii. Suture cart (velcro/sticky tag to side)
  - iii. Scanner (velcro/sticky tag to side)
  - iv. Monitor car (velcro/sticky tag to side)
  - v. Instrument table (velcro/sticky tag to side)



by vendor's

- d. Place one human near instrument table with tag on hip
  - e. Place one human near gurney with tag on hip
- 2. **Execution:**
  - a. Record the tags reported location for 30 seconds at the system's highest refresh rate
- 2. **Expected result:**
  - a. Reported location data for each tagged item
- 3. **Evaluation:**
  - a. This test will be evaluated based on the following criteria:
  - b. Accuracy (mean error and standard deviation)
  - c. Frequency of significantly erroneous data. (greater than twice the systems vendor proposed accuracy)
  - d. Frequency of "location jumps". (movement of more than half the systems vendor proposed accuracy)
- 4. **Reporting:**
  - a. Add to per vendor data for contrived OR test. This data will consist of:
  - b. Reported location data
  - c. Ground truth location data
  - d. Statistical analysis (mean error, standard deviation, etc)
  - e. Subjective evaluation of the above criteria
- 5. **Clean-up:**
  - a. Return resources to storage

**Test Number:** A2.1.1

**Test Name:** Dynamic location

**Author:** JMN

**Last Updated:** 10 July 2005

**Major Criterion:** Accuracy

**Minor Criterion:** Dynamic

**Estimated time for completion:** 2 hours

**Scenario:** Move a tag around OR A; track tag's true location with Vicon motion capture system.

**Resources needed:**

1. Tag
2. Measuring tape
3. Readers (Quantity: vendor-specified)
4. Large cabinets
5. Laptop with 802.11 b/g wireless NIC
6. Desktop PC with ample disk space, Gb NIC, CD writer
7. As-built engineering drawings of the test environment
8. Fully functional system
9. Misc. equipment required by vendor's system.
10. Vicon motion capture system
11. Dr. Lee
12. Person

**Materials consumed:**

1. Making tape

**Detailed steps:**

**1. Preparation:**

- a. Gather required resources
- b. Setup vendor's system
- c. Setup Vicon motion capture system
- d. Attach a reflector from the Vicon system to a tag with double sided tape
- e. Attach tag to belt clip with velcro/tape/string. Affix tag on belt, ensure clothing does not obstruct any part of the tag
- f. Measure a 4x4 meter test area; mark with duct tape

**2. Execution:**

- a. Walk across the test area, make a 90 degree turn, walk half a meter, and walk back across the test area until the entire area has been covered. Record the tag's reported location throughout this "walk" at the system's highest refresh rate
- b. Repeat the test at 1/2 and 2 times walking speed

**3. Expected result:**

- a. Reported location data for tag

- b.** True location data for tag (from Vicon system)
- 4. Evaluation:**
  - a.** None; this an objective measure.
- 5. Reporting:**
  - a.** Add to table of per vendor dynamic location results. These results will consist of:
    - i.** Table of reported location with timestamp from vendor's system
    - ii.** Table of reported location with timestamp from Vicon.
    - iii.** Percentage error of the system's reported location
    - iv.** Mean error and standard deviation
- 6. Clean-up:**
  - a.** Return resources to storage

**Test Number:** A3.1.1

**Test Name:** Co-location

**Author:** JMN

**Last Updated:** 11 July 2005

**Major Criterion:** Accuracy

**Minor Criterion:** Co-location

**Estimated time for completion:** 3 hours

**Scenario:** Test the system's ability to co-locate tags. Tags will be placed as close together as possible. Reported location data for each tag will be recorded as the tags are moved apart in pre-defined increments.

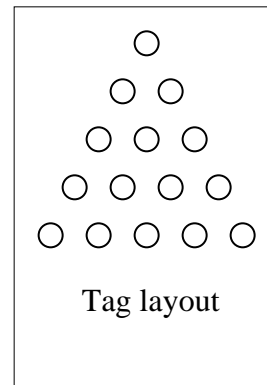
**Resources needed:**

1. Tags (Quantity: 15)
2. Measuring tape
3. Readers (Quantity: vendor-specified)
4. Large cabinets
5. Laptop with 802.11 b/g wireless NIC
6. Desktop PC with ample disk space, Gb NIC, CD writer
7. As-built engineering drawings of the test
8. Fully functional system
9. Misc. equipment required by vendor's system

environment

**Materials consumed:**

1. Making tape



**Detailed steps:**

**1. Preparation:**

- a. Gather required resources
- b. Setup vendor's system
- c. Place the 15 tags in an equilateral triangle as close as possible. Tags will be placed on a table at waist height

**2. Execution:**

- a. Record the tags reported locations for 30 seconds at the systems highest refresh rate
- b. Increase the distance between neighboring tags by 1 inch; repeat

**3. Expected result:**

- a. Reported location data for all 15 tags

**4. Evaluation:**

None; this is an objective measure

**5. Reporting:**

- a. Add to per vendor table of Co-location results. These results will consist of:
  - i. A table of reported locations and corresponding timestamps for each triangle size.
  - ii. A table of true locations and corresponding timestamps for each triangle size.

**6. Clean-up:**



- a.** Return resources to storage

**Test Number:** S2.1.1

**Test Name:** Integration effort (documentation)

**Author:** JMN

**Last Updated:** 12 July 2005

**Major Criterion:** System characteristics

**Minor Criterion:** Integration effort

**Estimated time for completion:** 2 hours

**Scenario:** Consult the system's documentation, manuals, API, etc to evaluate the effort involved in using the position data to drive another application. Consult a system representation to evaluate the effort involved in using the position data to drive another application.

**Resources needed:**

1. System manual/documentation/API
2. System representative

**Materials consumed:**

None

**Detailed steps:**

**1. Preparation:**

- a. Gather required resources

**2. Execution:**

- a. Evaluate the effort involved in using the position data to drive another application by reading the manual/documentation/API

**3. Expected result:**

- a. The effort required to use the position data to drive another application

**4. Evaluation:**

The system will be placed in one of five categories described below:

1. Closed system  
Position data cannot be accessed
2. Database access  
Position data is accessed via database queries
3. API event access  
Position data is accessed with an API or SDK.
4. Outbound messaging  
Position data can be sent to middleware.
5. Other

**5. Reporting:**

- a. Add to table of per vendor integration effort. The table will include:
  - i. A paragraph explaining the system's available methods of integration
  - ii. The system's placement into one of the four above categories

**6. Clean-up:**

- a. Return resources to storage

**Test Number:** S3.1.1

**Test Name:** RF interference

**Author:** JMN

**Last Updated:** July 12 2005

**Major Criterion:** System characteristics

**Minor Criterion:** RF characteristics

**Estimated time for completion:** 30 minutes

**Scenario:** Ascertain RF emissions of the system and compare them to RF emissions in a typical OR to identify possible conflicts.

**Resources needed:**

1. System specifications
2. List of RF equipment present in OR with the frequencies used

**Materials consumed:**

None

**Detailed steps:**

**1. Preparation:**

- a. Gather required resources

**2. Execution:**

- a. Compare the RF emissions of the system to the RF emissions of OR equipment to identify possible conflicts

**3. Expected result:**

- a. A list of equipment the system may conflict with

**4. Evaluation:**

None; this is an objective measure

**5. Reporting:**

- a. Add to table of per vendor RF conflicts. The table will include:
  - i. The devices the system may conflict with

**6. Clean-up:**

- a. Return resources to storage

**Test Number:** S4.1.1

**Test Name:** Interoperability

**Author:** JMN

**Last Updated:** 12 July 2005

**Major Criterion:** System characteristics

**Minor Criterion:** Interoperability

**Estimated time for completion:** 2 hours

**Scenario:** Evaluate the systems ability to operate in the same area as another vendor's system with concurrent static testing.

**Resources needed:**

1. All equipment supplied by vendor A to run A1.\*.\* tests
2. All equipment supplied by vendor B to run A1.\*.\* tests
3. Laptop with 802.11 b/g wireless NIC
4. Desktop PC with ample disk space, Gb NIC, CD writer
5. As-built engineering drawings of the test environment
6. Fully functional system
7. Misc. equipment required by vendors' system.

**Materials consumed:**

1. Blank CDs
2. Cat5 cable
3. Duct tape

**Detailed steps:**

**1. Preparation:**

- a. Gather required resources
- b. Setup system A
- c. Setup system B

**2. Execution:**

- a. Perform tests A1.\*.\* on system A
- b. Perform tests A1.\*.\* on system B

**3. Expected result:**

- a. Results of A1.\*.\* tests on system A and B

**4. Evaluation:**

None; this is an objective measure

**5. Reporting:**

- a. Add to table of per vendor interoperability test results. The test results be formatted by the A1.\*.\* specifications

**6. Clean-up:**

- a. Return resources to storage

**Test Number:** S5.1.1

**Test Name:** Privacy/Security

**Author:** JMN

**Last Updated:** 12 July 2005

**Major Criterion:** System characteristics

**Minor Criterion:** Privacy/Security

**Estimated time for completion:** 30 minutes

**Scenario:** Evaluate the privacy and security features of the system by consulting with the vendor and reviewing the documentation.

**Resources needed:**

1. System manual/documentation/API
2. System representative

**Materials consumed:**

None

**Detailed steps:**

**1. Preparation:**

- a. Gather required resources

**2. Execution:**

- a. Evaluate the privacy and security features of the system by reviewing the documentation.
- b. Evaluate the privacy and security features of the system by consulting with the vendor

**3. Expected result:**

- a. The privacy and security features of the system

**4. Evaluation:**

Systems will be evaluated on the following:

1. Ease of “sniffing” transmissions between system devices (i.e. running a network device in promiscuous mode)
2. Encryption of transmissions between system devices
3. Ease of physically “plugging into” a system device
4. Encryption of data stored on a system device
5. Ease of executing denial of service attack

**5. Reporting:**

- a. Add to table of per vendor privacy and security features
- b. Add to matrix of privacy and security features for each vendor

**6. Clean-up:**

- a. Return resources to storage

**Test Number:** S6.1.1

**Test Name:** Data communication load

**Author:** JMN

**Last Updated:** 12 July 2005

**Major Criterion:** System characteristics

**Minor Criterion:** Privacy/Security

**Estimated time for completion:** 30 minutes

**Scenario:** Assess that data communication load generated by the vendor's system.

**Resources needed:**

1. System representative
2. System documentation

**Materials consumed:**

None

**Detailed steps:**

**1. Preparation:**

- a. Gather required resources

**2. Execution:**

- a. Consult the system representative to determine the load generated by the vendor's system
- b. Consult the documentation to determine the load generated by the vendor's system

**3. Expected result:**

- a. Data communication load generated by the system

**4. Evaluation:**

None; this is an objective measure

**5. Reporting:**

- a. Add to per vendor results of data communication load test. This data will consist of the following:
  - i. Average transmission size in bytes
  - ii. Average interval of transmission
  - iii. Bytes an hour

**6. Clean-up:**

- a. Return resources to storage

**Test Number:** S7.1.1

**Test Name:** Battery strength

**Author:** JMN

**Last Updated:** 12 July 2005

**Major Criterion:** System characteristics

**Minor Criterion:** Sensitivity to battery strength

**Estimated time for completion:** 8 hours

**Scenario:** Drain batteries to various levels and re-perform static accuracy tests to the effect of battery strength on accuracy.

**Resources needed:**

1. All equipment supplied by vendor A to run A1.\*.\* tests
2. All equipment supplied by vendor B to run A1.\*.\* tests
3. Laptop with 802.11 b/g wireless NIC
4. Desktop PC with ample disk space, Gb NIC, CD writer
5. As-built engineering drawings of the test environment
6. Full functional system
7. Misc. equipment required by vendors' system.
8. System documentation

**Materials consumed:**

1. Masking tape

**Detailed steps:**

**1. Preparation:**

- a. Gather required resources
- b. Setup system A
- c. Drain all tag batteries to 75% power. The batteries will be drained using a resistive load. If they battery has a power alert function they battery power will be judged based on that report. Otherwise the percentage of power remaining will be estimated by:  $(1 - \text{amount of time resistive load applied} / \text{total battery life}) * 100$

**2. Execution:**

- a. Perform tests A1.\*.\* on system A
- b. Drain batteries to 50% power; repeat
- c. Drain batteries to 25% power; repeat

**3. Expected result:**

- a. Results of A1.\*.\* tests on systems A

**4. Evaluation**

- a. None; this is an objective measure

**5. Reporting:**

- a. Add to table of per vendor battery strength tests. The test results will be formatted by the A1.\*.\* specifications

**6. Clean-up:**

- a. Return resources to storage

**Test Number:** S8.1.1

**Test Name:** Tag data storage/transmission features

**Author:** JMN

**Last Updated:** 21 July 2005

**Major Criterion:** System characteristics

**Minor Criterion:** Tag data storage/transmission features

**Estimated time for completion:** 30 minutes

**Scenario:** Examine the system documentation to determine the data storage and transmission capabilities of the tag.

**Resources needed:**

1. System documentation

**Materials consumed:**

None

**Detailed steps:**

- 1. Preparation:**

- a. Gather required resources

- 2. Execution:**

- a. Examine documentation to determine tag data storage and transmission capabilities

- 3. Expected result:**

- a. Tag data storage capabilities in bits
  - b. Protocol used for transmitting tag data

- 4. Evaluation:**

None; this is an objective measure.

- 5. Reporting:**

- a. Add to per vendor data for tag data storage/transmission data results. This data will consist of:
    - i. Bits tag can store
    - ii. Protocol used for transmitting tag data

- 6. Clean-up:**

- a. Return resources to storage



**Test Number:** S9.1.1

**Test Name:** Bleed-through

**Author:** JMN

**Last Updated:** 3 August 2005

**Major Criterion:** System characteristics

**Minor Criterion:** Bleed-through

**Estimated time for completion:** 2 hours'

**Scenario:** Determine the bleed-through of the reader.

**Resources needed:**

1. Fully functional system
2. Laptop with 802.11 b/g wireless NIC
3. RF signal strength reader

**Materials consumed:**

None

**Detailed steps:**

**1. Preparation:**

- a. Gather required resources
- b. Setup system

**2. Execution:**

- a. Measure RF signal strength in areas adjacent to Sim center
- b. Measure RF signal strength above and below Sim center
- c. Measure 802.11b/g signal strength in areas adjacent to Sim center
- d. Measure 802.11b/g signal strength above and below Sim center

**3. Expected result:**

- a. RF signal strength in above areas
- b. 802.11b/g signal strength in above areas

**4. Evaluation:**

None; this is an objective measure.

**5. Reporting:**

- a. Add to per vendor data for tag bleed-through data results. This data will consist of:
  - i. RF signal strength in above areas
  - ii. 802.11b/g signal strength in above areas

**6. Clean-up:**

- a. Return resources to storage

## **Exhibit 1-C**

**DRAFT**

# Evaluation of Active RFID in the Perioperative Context\*

Duncan Clarke  
Fremont Associates, LLC  
Camden, South Carolina

November 30, 2005

## Abstract

During July through October of 2005 the Operating Room of the Future project at the University of Maryland Medical Center evaluated six active RFID systems. The evaluation consisted of hands-on testing of a variety of COTS systems employing the leading active RFID technologies—802.11 RF, proprietary RF, ultra-wideband, infrared and ultrasound.

Extensive testing demonstrated that no single system was able to address all of the cost and performance constraints necessary to support passive tracking of all aspects of perioperative workflow. Rather, it is recommended that the problem domain be subdivided based on the performance characteristics required by various sub-problems, and individual technologies be employed where their price/performance characteristics are most appropriate. The deployment of multiple technologies in a hierarchy of active RFID systems along side existing perioperative IT systems will require an emphasis on middleware to integrate disparate information sources accurately, safely and securely in real-time.

## 1 Introduction

In January of 2005 the Operating Room of the Future (ORF) project at the University of Maryland Medical Center (UMMC) initiated a year-long project to evaluate the applicability of commercial off-the-shelf (COTS) RFID technologies for tracking workflow in the perioperative context. The goal was to determine whether existing COTS solutions were suitable for creating infrastructure to support a multi-year research project studying advanced workflow management techniques in the perioperative environment. The overall objective is to increase patient throughput and patient, physician and staff satisfaction without increasing capital or operating expenses (*e.g.*, adding physical operating room space or increased staffing) or reducing perioperative safety.

The evaluation process had three major phases: (1) review of academic and trade literature on existing uses of RFID in medical applications; (2) review of RFID technologies and vendors; and (3) hands-on evaluation of key exemplars of each technology. The results of phases (1) and (2) were reported previously [1]. This document summarizes the findings of phase (3).

The remainder of this document is organized as follows. Section 2 presents the context within which the testing was performed, and the methodology employed during testing. Section 3 describes experimental results for each of the systems evaluated. Section 4 presents an analysis

---

\*This work was supported in part by USAMRMC-TATRC award DAMD 17-3-2-001.

of the experimental data, providing guidance for the application of COTS active RFID systems in the perioperative environment. Section 5 concludes the presentation with summary remarks. Appendices A-F provide extensive supporting data.

## 2 Methodology

In this section we describe the framework within which testing took place. This includes the project context in which judgments were made, the criteria for selecting candidate systems, and the test plan for evaluating systems.

### 2.1 Context

The UMMC ORF project seeks to apply computer-aided workflow management techniques to the perioperative process. Within the scope of this work we define *perioperative* broadly, to include potentially every activity from initial consultation through discharge of a surgical patient. The project seeks to apply workflow management in a way that will cost effectively increase OR work velocity. Other overriding concerns, common to all projects in the perioperative domain, are (i) patient, physician and staff safety, (ii) patient, physician and staff satisfaction; and (iii) security and privacy of clinical data.

It is worth noting that physician and staff satisfaction will be heavily influenced by the degree to which a new work process disrupts existing practices. Thus changes that can improve the flow of work and information within existing practices are preferred. Where disruptive change is necessary costs of planning and training for transition must be accounted for in the initial cost/benefit analysis.

IT-based process changes typically involve introducing traditional PC workstations (stationary or portable) and context-intensive application programs that require uninterrupted interaction. Both PC workstations and context-intensive applications have proven difficult to integrate into the perioperative workplace [4, 3, 2] due to their requirement for focus on a single task at a dedicated workstation. To avoid these well known problems, our project has focused on technologies for passive tracking of people, objects, documents, *etc.*, specifically, RFID [10].

RFID systems can be coarsely divided into two basic types—*passive* RFID [9] which uses powered radio transceivers to communicate with battery-less radio “tags” which reply by means of harvested RF power, and *active* RFID which augments passive RFID tags with built in batteries to autonomously power computation and radio reception and transmission.

Passive RFID systems have received much attention in the popular press due to big-box retailer requirements for their use in supply chain management [8, 7]. Though we can envision the application of passive RFID in perioperative workflow management, we have chosen not to analyze and report on these technologies within the scope of this project. A companion project at the UMMC has already undertaken a detailed investigation of passive RFID technologies and we intend to leverage their work in selecting such systems for application where appropriate.

At the outset of the evaluation process a list of potential tracking objectives were identified:

- Patient Tracking—determine the location of patients in real-time.
- Asset Allocation Tracking—determine on a periodic basis which assets (*e.g.*, beds, infusion pumps, *etc.*) are in active use.

- Asset Tracking—locate critical capital assets in real-time to improve utilization and reduce the risk of theft.
- Staff Tracking—report the location of physicians and staff in real-time.
- Document Tracking—report the location of critical documents (*e.g.*, patient consent for surgery), the absence of which can impede workflow.
- Instrumentation and Supply Tracking—track instrumentation and supplies on a per-case basis to improve surgical workflow, and on a pure device or supply category to track maintenance and/or improve utilization.
- Event Detection Based on Collocation—determine the occurrence (or non-occurrence) of significant events based on collocation of patients, physicians, staff, equipment, *etc.* [6].

With these applications in mind we set about selecting potential COTS active RFID systems for deployment across the perioperative environment.

## 2.2 Candidate Selection

The market for COTS active RFID systems is in its infancy. Consequently there is no dominant vendor, and many small vendors compete with systems based on disparate technologies and proprietary or, at best, non-standard interfaces. From the long list of potential vendors we selected at least one vendor for each of the important technologies, based on market presence, interaction with their sales and technical staff, and price.

The vendors whose systems were finally purchased and evaluated were as follows (in no particular order):

### 2.2.1 Ekahau.

A software-based system that uses existing 802.11 access points to track any 802.11 devices. Ekahau sells application-specific devices (tags) for tracking staff or equipment; these tags were used during the evaluation process. Figure 3 illustrates the Ekahau tag.

Position is reported in terms of (X,Y) coordinates on the floor plan of the operating environment. If zones (*i.e.*, rooms, units, floors, *etc.*) are defined the system can also report location by zone. Figure 2 illustrates the Ekahau software application interface.

### 2.2.2 Sonitor.

Hardware and software that uses ultrasound to track proprietary tags. Figure 16 illustrates the Sonitor tag. Figure 17 illustrates the Sonitor reader.

Position is reported in terms of the nearest reader location. If reader locations are correlated to zones the system will report location by zone. Figure 15 illustrates the Sonitor software application interface.

### 2.2.3 Radianse.

Hardware and software that uses a combination of proprietary RF and infra-red light (IR) to track proprietary tags. Figure 24 illustrates the Radianse tag. Figure 25 illustrates the Radianse reader.

Position is reported in terms of pre-defined zones. Figure 23 illustrates the Radianse software application interface.

### 2.2.4 UbiSense.

Hardware and software that uses ultra-wideband RF signals and highly directional receivers to track proprietary tags. Figure 31 illustrates the UbiSense tag. Figure 32 illustrates the UbiSense reader.

Position is reported in terms of (X,Y,Z) coordinates in the space being monitored.

### 2.2.5 Agility Healthcare.

Hardware and software that uses proprietary RF at two different frequencies to track proprietary tags. Multiple reader types to monitor zones and portals. Figure 46 illustrates an assortment of Agility tags. Figures 47 and 48 illustrate the Agility readers.

Position is reported in terms of pre-defined zones. Figure 45 illustrates a simple position tracking software application's interface. Agility's business focus is on the larger issue of asset management, so this image is not representative of their flagship application.

### 2.2.6 Exavera.

Hardware and software that uses proprietary RF to track proprietary tags. A novel combination of wired base-stations and low-cost wireless "relays" is used to improve coverage without the need for large numbers of full-function transceivers. Figures 54 through 56 illustrate the Exavera tags. Figures 57 and 58 illustrate the Exavera reader and relays.

Position is reported in terms of (X,Y) coordinates and pre-defined zones.

The primary application for which all of the above systems are currently being marketed is asset, staff and patient tracking within a hospital or medical center. The UbiSense system is marketed more generally as a high precision tracking solution for smaller spaces. Our testing focused on their suitability for use in the applications listed in Section 2.1. Toward that end we focused on issues of suitability for the perioperative milieu, and accuracy and precision of position reporting.

## 2.3 Test Plan

In August of 2005 we created a comprehensive test plan based on objective criteria that we believed were critical to the perioperative environment [5]. This included basic tests of static and dynamic position reporting accuracy, as well as more domain-specific tests such as tolerance for chemical and heat sterilization. It quickly became apparent that current COTS systems do not deliver the positional accuracy or tolerate sterilization in ways that would produce meaningful results for many of these tests.

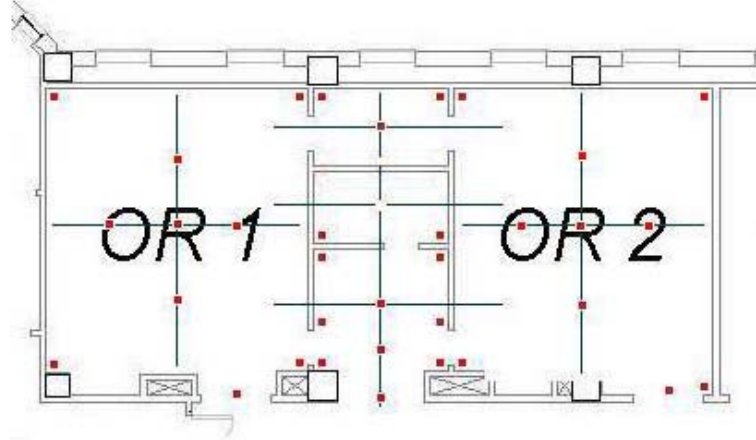


Figure 1: Physical Test Environment

The work reported here focused on a subset of tests from the original test plan. These tests were chosen to be challenging, yet produce meaningful results for all systems. To insure comparability of results, we had each vendor provide a system appropriate for a standard environmental configuration. The one exception to this standard configuration profile was the UbiSense system, where cost prohibited our instrumenting the entire test area with sensors.

Identical tests of location reporting abilities and analytic tests of the systems and their components were performed, as described in the following.

### 2.3.1 Configuration

Each vendor was given a AutoCAD file containing a floor plan identical to that shown in Figure 1. It depicts two decommissioned operating rooms and three wet lab/storage spaces in between. The vendors were told that we wanted to record the location of assets in the OR's and the spaces in between with the best accuracy and precision their system would allow. The quantity and placement of readers was left to the vendor.

Vendors were asked to supply 15 tags of assorted types. The following exceptions are noted: (1) Sonitor supplied 15 tags, but one was damaged during testing; (2) Agility supplied 11 tags (due to miscommunication between sales and technical staff), one defective (cause unknown); (3) UbiSense supplied five tags (due to cost constraints).

Each vendor delivered and configured their system and performed whatever tuning was necessary. Set-up, configuration and test was performed by vendor technicians on site in the test environment at UMMC. Immediately after the vendor's representatives left the system was tested using the physical and software configuration provided by the vendor.

Each system was tested independently. No attempt was made to evaluate the interoperability of systems from multiple vendors.

The 802.11 based system (Ekahau) was set up by vendor representatives using the same Enterasys RoamAbout access points currently in use at UMMC. We provided the vendor with as many access points as they required to create a stand-alone 802.11 network specifically for the purpose of tracking tags in our test environment.

The 802.11/802.3 networks used by the systems under test were stand-alone IP networks. They were not connected to the hospital network or the larger Internet in any way. A desktop PC running Windows XP was provided for use as a server, or vendors could bring their own host hardware.

The decommissioned OR spaces were still populated with lights, tables, cabinets, *etc.* like you would find in working OR's. Walls were covered with ceramic tile from floor to ceiling. All portals shown in Figure 1 had working doors that were controlled during the testing.

### 2.3.2 Static Tests

Static tests involved placing one or more tags at each of the test points indicated on Figure 1. These test points were not revealed to the vendors in advance of the tests or during their on-site configuration. Test points varied in their height and structure. Some points were on top of stainless steel counters; some points were mid-floor, where the tag was placed on a cardboard stand to keep it above floor level; some points were on stainless steel shelves very close to floor level; some points were on cast iron and porcelain scrub sinks; some points were on steel window sills.

After placing the tag, the test engineer returned to the central workstation (located in OR 2), waited at least two sample periods for the system to stabilize, and recorded the location data reported by the system. Tests that involved simple location of objects used a single tag, with other tags powered down or moved sufficiently far from the test environment that they could not interfere with the readers. Some tests required the use of multiple tags simultaneously.

The static test scenarios are as follows:

S01—Doors Closed. All doors in the test environment were closed and a single tag was moved from location to location.

S02—Doors Open. All doors in the test environment were opened and a single tag was moved from location to location. Since doors can attenuate some signals, the configuration of doors was considered a significant variable in the test process.

S03—Doors Closed, Fabric Cover. A single tag was placed in the breast pocket of a scrub top. The scrub top, containing the tag, was moved from location to location. The scrub top was intended to simulate the effect of a simple clothing cover that might be encountered in the tagging of people.

S04—Doors Open, Fabric Cover. Identical to scenario S03, except all doors were opened.

S05—Saline Bag, Exposed. A single tag was attached to the top face of a three liter bag of saline. This could either be viewed as (1) a test of the tagging of a particular low-value medical supply; or (2) a test of the tagging of humans, who have many of the same radio reflection/attenuation properties as a bag of water. The saline bag, with the attached tag, was moved from location to location. Care was taken to insure that the tag was on an exposed face at all times.



- S06—Saline Bag, Concealed. Identical to scenario S05, except care was taken to insure that the tag was on an unexposed face at all times. That is, the saline bag covered the tag, likely interfering with its signal(s).
- S07—Instrument Tray, Outside. A single tag was attached to the exterior of a tray of instruments intended for open abdominal surgery. The instrument case itself was made of medical-grade aluminum, a large volume of steel instruments was contained inside. The instrument case was moved from test point to test point.
- S08—Instrument Tray, Inside. A single tag was placed inside the instrument case, and the instrument case was closed. Note that though the instrument case is made of metal, there are vents in the case top intended to allow gasses to enter and exit the case during sterilization. If a signal could escape the case, testing proceeded with the instrument case moved from test point to test point. Note that in most cases no data was received in this configuration.
- S09—Gurney. A single tag was attached to the undercarriage of a standard hospital gurney. The gurney was moved from test point to test point, for test points that allowed placement of the gurney.
- S10—Repeatability. A single tag was placed at the center of OR 1, a test point recorded, the tag was moved to the center of OR 2, a test point was recorded, and this process was repeated for 15 iterations. This test was intended to evaluate the repeatability of results from sample to sample. In theory the systems should have reported the same location data on each iteration. In practice, the accuracy bound of each system determined the amount of noise that was present in the set of samples.
- S11—High Volume, No Line of Sight. All available tags were placed in a cardboard box and shuttled between OR's one and two for two iterations, as in scenario S10. A systems ability to manage multiple reporting tags in a small area will be critical if tagging is to be ubiquitous.
- S12—Collocation. Twelve tags (five in the case of UbiSense) were placed in the center of the OR in a circular pattern. The radius of the circle was increased in half-meter steps from zero to two meters. The position of each tag was recorded. This process was carried out once in each OR. This test evaluated each system's ability to discern whether tags were co-located (close in linear space, or located in the same zone).

Figures 70 through 80 illustrate these scenarios.

### 2.3.3 Dynamic Tests

All plans for dynamic tests (*i.e.*, sampling tag position while tags are in motion) were canceled once the accuracy and refresh rates of COTS systems were understood. The UbiSense system could have been used to produce significant results in this area, but there was no other system that would have served as a meaningful point of comparison.

### 2.3.4 Analytic Tests

Analytic tests involved evaluating the physical characteristics of tags and readers, and characteristics of the software systems such as technology employed, ease of set-up and use, cost, ease of integration with other systems, *etc.*

## 3 Experimental Results

The tests outlined in Section 2 were performed over a three month period from August through October of 2005. Results were recorded in tabular form and stored in Excel spreadsheets. Those results were summarized in graphical representations included in the appendices. In this section we will summarize our experience with and results measured for each system.

### 3.1 Ekahau

The total cost of the Ekahau system including equipment and labor for on-site set-up was \$10,895. Part of this cost (\$1,500) covered Ekahau T-201 wireless location tags, which were not absolutely required for system operation. The cost of the Enterasys RoamAbout access points used is not included in this figure.

The Ekahau system computes an (X,Y) position from received signal strength, and reports this position as well as an inferred zone identifier based on pre-defined zones.

Results of static testing (Figures 4 through 7) showed median location error on the order of three meters. This is considered a typical level of accuracy for 802.11 based location systems. The accuracy of inferred zone data was relatively poor because the system has no facility for differentiating zones other than (X,Y) position, and this position data was not sufficiently accurate to differentiate the zones in our test environment. No signal was received from the Ekahau tag when it was placed inside the closed instrument tray.

Results of repeatability testing (Figures 8 through 10) were similar, on a point-wise basis, to the previous static tests. The zone-level accuracy was much higher due to the location of the test points for this scenario (in the center of the large OR's).

Results of high volume, no line of sight testing (Figures 11 through 13) were degraded from the repeatability testing results, likely due to RF signal interference from the large number of tags reporting from a confined area.

Results of collocation testing (Figure 14) were in line with other static tests.

Though any 802.11 device can be tracked, we specifically evaluated the Ekahau T-201 tags. These tags were relatively small, at  $5.5 \times 5.0 \times 2.9$  cm. Tags (including battery) weigh 77.5 g. Tags are equipped with a removable belt clip for use in person-tagging applications. Each tag has an integrated rechargeable battery that is recharged through an external plug.

### 3.2 Sonitor

The total cost of the Sonitor system including equipment and labor for on-site set-up was \$10,928.

The Sonitor system computes a nearest-sensor position from received signal strength, and reports an inferred zone identifier based on pre-defined zones. The system does not compute or report specific (X,Y) coordinates.

Results of static testing (Figures 18 and 19) indicated only occasional zone inaccuracy when doors were closed separating rooms in the test environment. When doors were opened, allowing sound to travel more easily between zones, accurate zone reporting became more problematic.

The Sonitor tags use an integrated motion sensor to determine when to broadcast a signal. The tags do not receive communication from any central controller, nor do they broadcast at periodic intervals. The goal of this design was to maximize battery life. However, the test tag used during our evaluation had a motion sensor that was not very sensitive in a horizontal orientation. Tests that involved anchoring the tag to massive objects (*i.e.*, the instrument tray and gurney) often failed to activate the tag's motion sensor. Thus, limited data was available for the instrument tray (outside) case, and no data was available for the instrument tray (inside) and gurney cases.

Results of repeatability testing (Figure 20) showed high accuracy when tags were positioned in the center of the OR with the doors closed.

Results of high volume, no line of sight testing (Figure 21) showed high accuracy when a high volume of tags was positioned in the center of the OR with the doors closed.

Results of collocation testing (Figure 22) showed high accuracy when a tags were relatively close to the center of the room, but accuracy dropped as tags approached boundary areas with other zones.

The provided Sonitor tags were a late-stage prototype intended to be worn in the shirt pocket, like a pen. Size was  $11.5 \times 1.7 \times 2.7$  cm. Tags (including battery) weigh 34.73 g. Each tag has a replaceable, disposable battery similar in size to standard AA batteries. The prototypes we were provided had the batteries soldered directly to leads, but production versions will have standard battery clips.

### 3.3 Radianse

The total cost of the Radianse system including equipment and labor for on-site set-up was \$16,427. Radianse actually delivered hardware sufficient to cover the entire OR wing, which included two more operating rooms, a long corridor and assorted office and lab/storage spaces.

The Radianse system computes a zone-level position from received RF and IR signal strength. The system does not report specific (X,Y) coordinates.

Results of static testing (Figures 26 and 27) showed low zonal accuracy overall, though the errors were almost exclusively in the reporting of tags located in the lab/storage areas. It appears that the system configuration provided did not adequately differentiate these small areas to provide meaningful results. No signal was received from the Radianse tag when it was placed inside the closed instrument tray.

Results of repeatability testing (Figure 28) showed anomalous behavior for tags entering OR 2, with one third of samples reporting incorrectly that the tag was in a lab/storage area.

Results of high volume, no line of sight testing (Figure 29) showed high accuracy. Note that these tests were carried out exclusively in OR 1 and OR 2.

Results of collocation testing (Figure 30) showed showed high accuracy. Note that these tests were carried out exclusively in OR 1 and OR 2 with minimal movement between zones.

Radianse tags are designed to be affixed to equipment, or clipped to a person's clothing. Size was  $4.5 \times 7.3 \times 1.9$  cm. Tags (including battery) weigh 35.87 g. Each tag has a replaceable, disposable "button" battery.

### 3.4 UbiSense

The total cost of the UbiSense system including equipment and labor for on-site set-up was \$19,950. This was for a non-standard configuration that covered only OR 2, and included only five tags.

This system was fundamentally different from all other systems evaluated in terms of technology, accuracy, complexity of configuration and cost. Reported data is very high accuracy (X,Y,Z) coordinates within the monitored zone. The accuracy attained in our configuration was limited by the temporary mounting of sensors used in all scenarios. Sensors were mounted relatively low in the room, and one or more meters from the corners, limiting their ability to cover the corners of the room where several test points resided.

The system configuration required a standard 802.3 network cables for each sensor, as well as additional "timing" cabling between readers. The high connection complexity of this system would be a critical consideration in a large deployment.

Results of static testing (Figures 33 through 37) showed location reporting error in three dimensions consistently below three meters, with median values in the tens of centimeters. In the area of the room well covered by the sensors location reporting errors were below ten centimeters. This is a two order of magnitude improvement over the other RF-based systems evaluated. No signal was received from the UbiSense tag when it was concealed under the saline bag or placed inside the closed instrument tray.

Results of repeatability testing (Figures 38 through 40) showed very high accuracy for this mid-zone scenario.

Results of high volume, no line of sight testing (Figures 41 through 43) showed very high accuracy for this mid-zone scenario, though error levels did increase over the previous scenario. Note also that the volume of tags was not as high as in other tests, since only five UbiSense tags were available.

Results of collocation testing (Figure 44) showed very high accuracy for this mid-zone scenario.

UbiSense tags are designed to be affixed to equipment, clipped to a person's clothing, or worn on a lanyard. Size was  $9.3 \times 5.9 \times 1.6$  cm. Tags (including battery) weigh 56.5 g. Each tag has a replaceable, disposable "button" battery.

### 3.5 Agility Healthcare

The total cost of the Agility Healthcare system including equipment and labor for on-site set-up was \$13,400.

The Agility Healthcare system computes a zone-level position from received RF signal strength. In order to specifically address zone-level accuracy, the system uses portal controllers that activate a tag as it passes through a designated portal so that it can report its movement. The system does not report specific (X,Y) coordinates.

Results of static testing (Figures 49 and 50) showed relatively poor zone-level accuracy, especially in the lab/storage areas. Interestingly, a signal was received from the Agility tag when it was placed inside the closed instrument tray. Results were comparable to other, less challenging cases.

Results of repeatability testing (Figure 51) showed relatively poor zone-level accuracy due to anomalous behavior as the tag passed from OR 1 back to OR 2.

Results of high volume, no line of sight testing (Figure 52) showed relatively poor zone-level accuracy on the first iteration, perfect results on the second.

Results of collocation testing (Figure 53) again showed moderate zone-level accuracy in OR 1, poor accuracy in OR 2.

The overall poor showing at zone-level accuracy is surprising, given that the system seems to have been engineered specifically for accurately capturing movement between zones. This is likely due to the fact that miscommunication within Agility Healthcare resulted in technicians arriving on site with a configuration designed for an environment different from that outlined in Agility's final proposal (which accurately reflected the test environment). Technicians adapted the system to the actual test environment as best they could while on site. Resource constraints at UMMC did not allow for re-testing with a fully redesigned configuration, so the accuracy of a properly provisioned Agility system remains an open question.

Agility Healthcare provides a range of different tags designed to be affixed to equipment, clipped to a person's clothing, or worn as a bracelet. The asset tag's size was a tiny  $2.4 \times 4.8 \times 0.8$  cm. Asset tags (including battery) weigh 8.3 g. The wrist tag's size was again quite small, at  $2.5 \times 3.8 \times 1.1$  cm. Wrist tags (including battery) weigh 7.0 g. Asset tags have a non-replaceable integrated battery. Wrist tags have a replaceable, disposable "button" battery.

### 3.6 Exavera

The total cost of the Exavera system including equipment and labor for on-site set-up was \$4,050.

The Exavera system computes an (X,Y) position from received signal strength, and reports this position as well as an inferred zone identifier based on pre-defined zones.

Results of static testing (Figures 59 through 62) showed median location error on the order of two to three meters. The Exavera system has improved slightly on the accuracy achieved by pure 802.11 based systems by employing a higher density of low cost receive-only relay devices.

The accuracy of inferred zone data was relatively poor because the system has no facility for differentiating zones other than (X,Y) position, and this position data was not sufficiently accurate to differentiate the zones in our test environment. No signal was received from the Exavera tag when it was placed inside the closed instrument tray. As was typically the case for all systems, the most difficult zones to resolve were the small lab/storage areas between the OR's. No signal was received from the Exavera tag when it was placed inside the closed instrument tray.

Results of repeatability testing (Figures 63 through 65) showed high accuracy for these tests placing tags in the center of the two OR's. There was no "memory" effect (*i.e.*, remembering a position in an intermediate zone) in transitioning from test to test.

Results of high volume, no line of sight testing (Figures 66 through 68) were similar, though position errors were rising to levels that resulted in missed zone assignments.

Results of collocation testing (Figure 69) showed excellent results for OR 1. Results in OR 2 were more problematic.

The tags provided by Exavera were early manufacturing prototypes, not representative of production versions currently in development. Thus meaningful dimensional and weight information was not available at the time of testing.

## 4 Analysis

### 4.1 Accuracy

The systems evaluated are primarily marketed for asset tracking applications within large hospitals or medical centers. For that application, they provide a reasonable approximation of location. A staff member in search of a specific piece of equipment or fellow staff member could easily find the object or person in question when given location data accurate to within three meters. Human actors are very adaptable and forgiving in these sorts of tasks.

On the other hand, overall accuracy results were significantly below any kind of reasonable standard for safety-critical perioperative systems. The only system approaching a level of accuracy that might allow differentiation of individual persons or objects within a room was the UbiSense system. However, the cost of the UbiSense hardware prohibits wide scale deployment.

Note also that these results were achieved in a very controlled environment where major pieces of equipment were stationary, and only one person was moving through the test environment. Adding the dynamic nature of the working OR's will undoubtedly negatively impact accuracy of positional and/or zonal data.

Another key consideration that was not touched on in this evaluation was the ability of systems to scale up for hospital-wide deployment. Though the cost limits imposed by the UbiSense system were pointed out, there are other technical considerations to be taken into account with the other systems. Specifically, does the provided software for managing sensors and tags scale to the level needed? Can the system differentiate objects on different levels in all scenarios that will arise, including multi-story atrium areas? Can the system differentiate adequately between logically unrelated areas that are physically adjacent?

### 4.2 Critical Considerations

Throughout the analysis tag power (*i.e.*, batteries) appeared as a critical design consideration and performance limiting factor. Battery size determines tag usage cycles (for rechargeable and replaceable batteries) and tag life (for disposable tags). Larger batteries reduce lengthen maintenance intervals or tag life. However, battery size also determines tag form factor. Batteries are the key determining factor in tag size and weight.

Because battery life is a critical design consideration, the systems evaluated employ a wide range of battery life management techniques. Some tags use motion sensors to limit transmission to times when the tag is detected to be in motion. Exclusive reliance on motion sensors is problematic if a tag is lost in a location where it is unlikely to be disturbed, or motion is insufficient to activate the tag's motion sensor. Tags employing regular, periodic broadcasts reduce the complexity of tag design (no motion sensor is required) and reduce the risk of tag loss, but at the expense of a predictable drain on the tag battery.

The Agility Healthcare system employs portal transceivers as a third way around the battery management problem. Since detecting absolute motion is not critical in systems designed for zone-level accuracy (only passage between zones is relevant), their tags use portal controllers to excite tags as they transition between zones. Thus broadcasts are produced only when needed.

Battery life will ultimately be a key factor in determining system life cycle costs. Systems with rechargeable batteries will require regular harvesting of tags for recharging and redeployment once charged. Systems with replaceable batteries will require regular monitoring of battery level, and

servicing of tags before a dead battery takes them out of service. Systems with disposable tags will require regular monitoring of battery level, and replacement of entire tags before a dead battery takes them out of service.

Contrast the need for vigilance with regard to tag batteries with the encode-and-forget nature of current technologies like bar codes and magnetic stripe cards, as well as passive RFID tags. When discussing technologies to enable passive tracking of objects, it is important to evaluate just how "passive" the new technology infrastructure is going to be from a deployment and maintenance stand point.

### 4.3 Suitability for Perioperative Workflow Management

Any application of these COTS systems to the automation of perioperative workflow management should be undertaken with great care. The accuracy of systems available at a price point allowing broad deployment is problematic for safety critical applications.

Even workplace decision making that is typically not thought of as safety critical, such as billing decisions, cannot rely on systems with the current level of accuracy found in COTS systems. For example, if automated workflow management determines that patient X is in the vicinity of care giver Y and object Z, it would be interesting to conclude that procedure P must have been initiated. However, putting that conclusion into action as an automatically generated billing event will require higher accuracy than is currently available.

Given these considerations, COTS active RFID systems are clearly not the entire solution to automation of perioperative workflow. Such applications are going to require a more sophisticated, hierarchical approach that integrates data from multiple sources of quantifiable accuracy.

Hospital wide, low cost systems can be used to provide a first-order estimate of location for a wide variety of patients, staff and resources. As the focus narrows to decision points more accurate systems will be employed within these narrower physical contexts. When high levels of accuracy is required systems that provide positive confirmation of events or presence, such as bar codes, swipe cards or biometric data will have to be employed.

The critical element in integrating these disparate technologies into a single system will be middleware that gathers information and its accompanying meta-information (source, accuracy, past reliability, *etc.*) and employs risk-aware decision making at all times.

## 5 Summary

The foregoing evaluation has focused on a narrow question—the technical suitability of existing COTS active RFID systems for application to perioperative workflow automation. The conclusions are mixed—there are relatively low cost systems that can produce position data that is usually accurate to the room level for a wide range of equipment and staff. However there are important operational costs that are not well understood, as well as operational issues that will likely impact system accuracy as the workplace evolves.

It should be noted that all of the systems evaluated are works in progress. All tags, readers and applications are evolving as the market matures and the focus narrows to critical applications. Thus, this evaluation carried out in the fall of 2005 is aging rapidly as vendors introduce smaller, lighter next-generation tags and more feature-rich versions of application software that is still in its infancy.

The current emphasis on hospital-wide asset tracking has resulted in systems that emphasize cost (capital costs and ongoing maintenance costs) at the expense of accuracy. This is a trend that should be expected to continue for several years, until the asset tracking market is saturated and vendors begin to target more demanding operational areas of the medical center.

Not touched on at all in this study were critical practical factors such as social acceptance of passive tracking of persons, or objects that could be casually linked to individual persons. Regardless of the technical attributes of the underlying technology, successful deployment in a social environment will hinge on social considerations.

## References

- [1] B. Bame, D. Clarke, M. Connolly, M. Minear, and A. Park. A survey of location technologies and their application in health care. Technical report, University of Maryland Medical Center, April 2005.
- [2] J. E. Bardram. Applications of context-aware computing in hospital work: examples and design principles. In H. M. Haddad, G. Papadopoulos, and A. Andreou, editors, *Proceedings of the 2004 ACM Symposium on Applied Computing*, pages 1574–1579. ACM Press, March 2004.
- [3] C. Bossen. The parameters of common information spaces: The heterogeneity of cooperative work at a hospital ward. In *Proceedings of the 2002 ACM conference on Computer Supported Cooperative Work*, pages 176–185. ACM Press, 2002.
- [4] H. B. Christensen and J. E. Bardram. Supporting human activities—exploring activity-centred computing. In G. Borriello and L. E. Holmquist, editors, *Proceedings of Ubicomp 2002: Ubiquitous Computing*, pages 107–116. Springer Verlag, September 2002. LNCS 2498.
- [5] J. Naslund, B. Bame, and D. Clarke. Ummc indoor positionin system evaluation facility. Technical report, University of Maryland Medical Center, August 2005.
- [6] Warren S. Sandberg, Matti Hkkinen, Marie Egan, Paige K. Curran, Pamela Fairbrother, Ken Choquette, Bethany Daily, Jukka-Pekka Sarkka, and David Rattner. Automatic detection and notification of "wrong patientwrong location" errors in the operating room. *Surgical Innovation*, 12(3):253–260, September 2005.
- [7] C. Silwa. Wal-mart supplier shoulder burden of daunting RFID effort. *Computerworld*, November 2003.
- [8] C. Silwa. Target issues RFID mandate to suppliers. *Computerworld*, March 2004.
- [9] H. Stockman. Communicating by means of reflected power. *Proceedings of the Institute of Radio Engineers*, pages 1196–1204, October 1948.
- [10] R. Want. RFID: A key to automating everything. *Scientific American*, pages 56–65, January 2004.



## A Ekahau

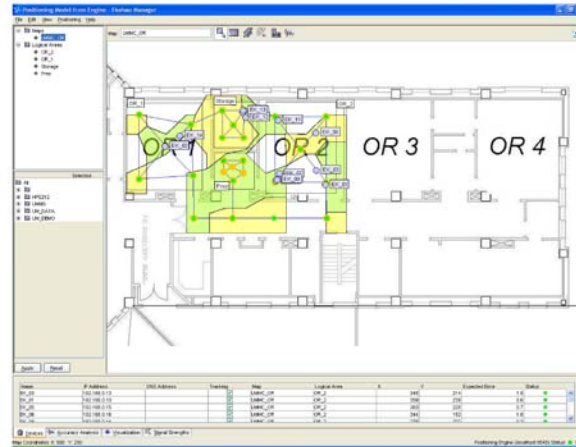


Figure 2: Application Screen, Ekahau



Figure 3: Tag, Ekahau

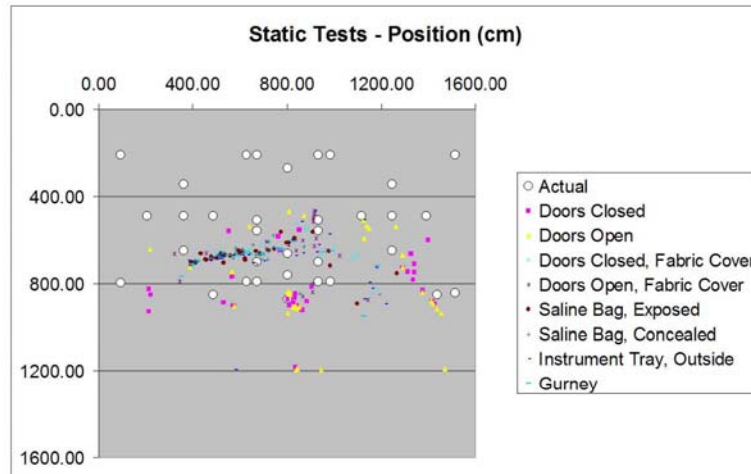


Figure 4: Single Tag Tests; Reported Position vs. Actual, Ekahau

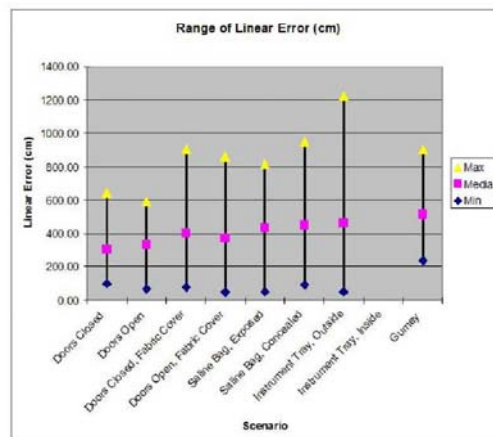


Figure 5: Single Tag Tests; Range of Error in Reported Position, Ekahau

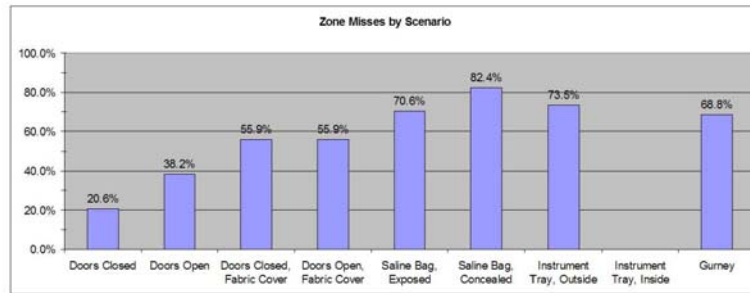


Figure 6: Single Tag Tests; Reported Zone vs. Actual, Ekahau

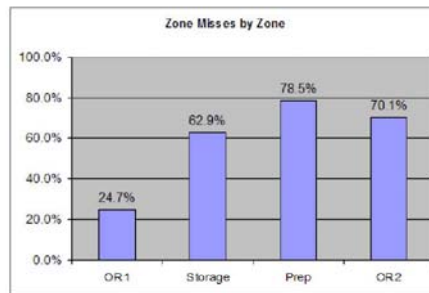


Figure 7: Single Tag Tests; Reported Zone vs. Actual, Ekahau

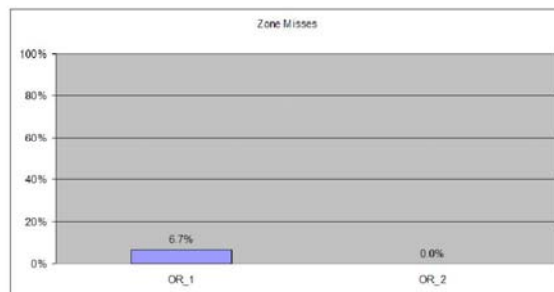


Figure 8: Repeatability Tests; Reported Zone vs. Actual, Ekahau

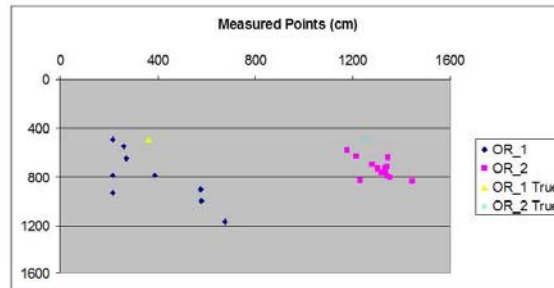


Figure 9: Repeatability Tests; Reported Position vs. Actual, Ekahau

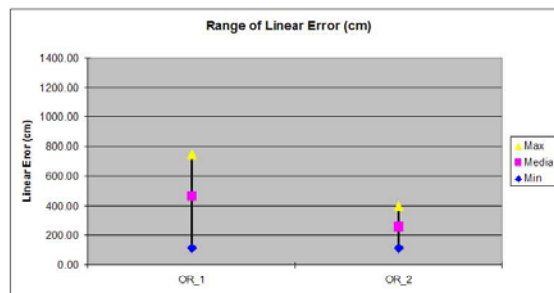


Figure 10: Repeatability Tests; Range of Error in Reported Position, Ekahau

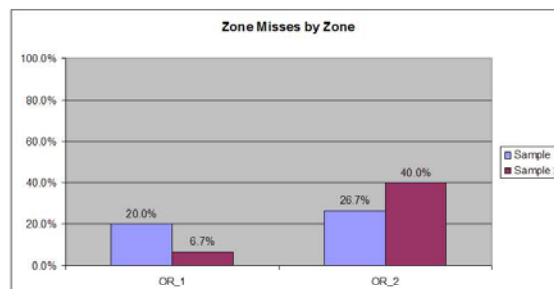


Figure 11: High Volume Tests; Reported Zone vs. Actual, Ekahau

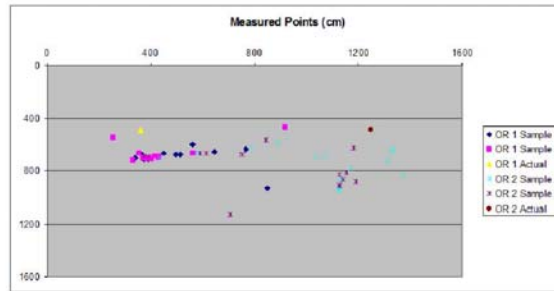


Figure 12: High Volume Tests; Reported Position vs. Actual, Ekahau

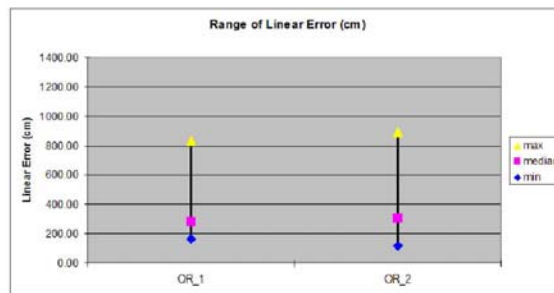


Figure 13: High Volume Tests; Range of Error in Reported Position, Ekahau

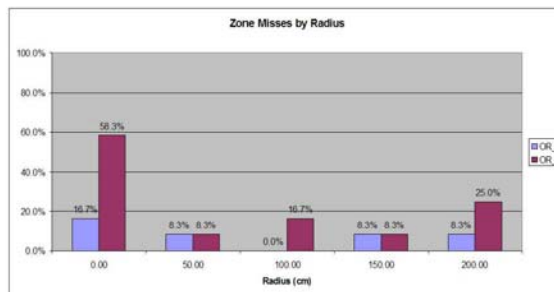


Figure 14: Colocation Tests; Reported Zone vs. Actual, Ekahau

## B Sonitor

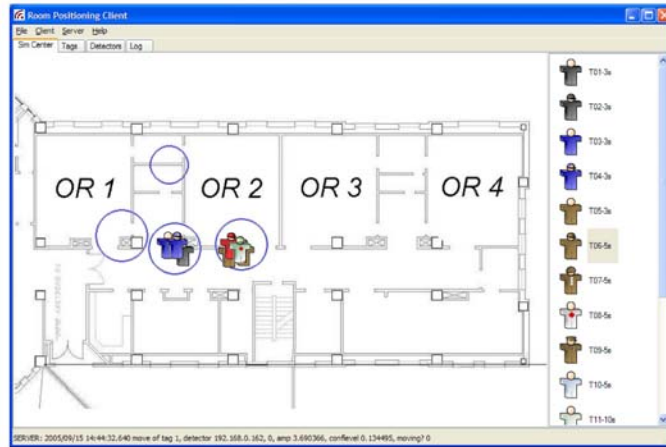


Figure 15: Application Screen, Sonitor



Figure 16: Tag, Sonitor



Figure 17: Reader, Sonitor

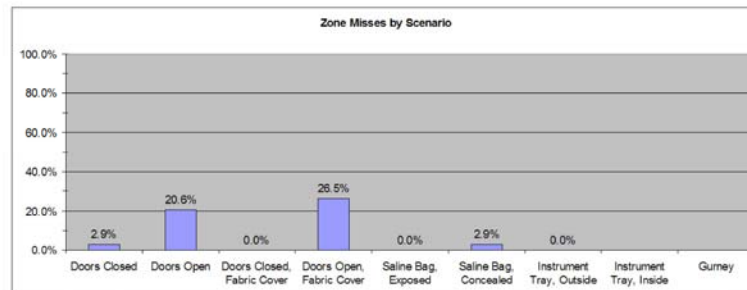


Figure 18: Single Tag Tests; Reported Zone vs. Actual, Sonitor

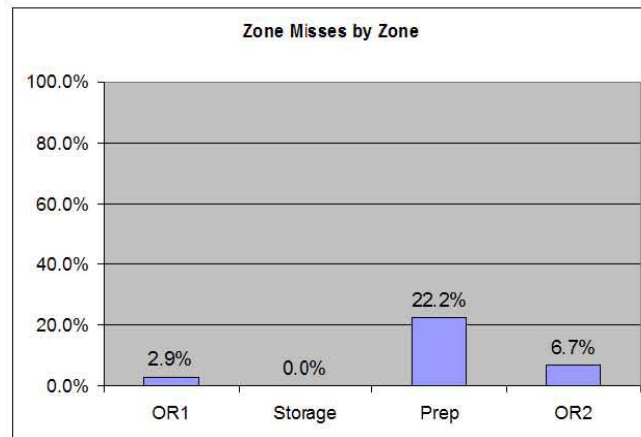


Figure 19: Single Tag Tests; Reported Zone vs. Actual, Sonitor

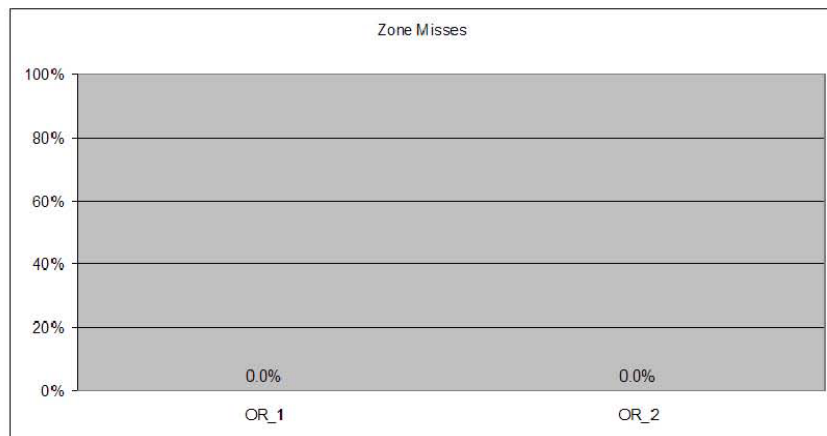


Figure 20: Repeatability Tests; Reported Zone vs. Actual, Sonitor



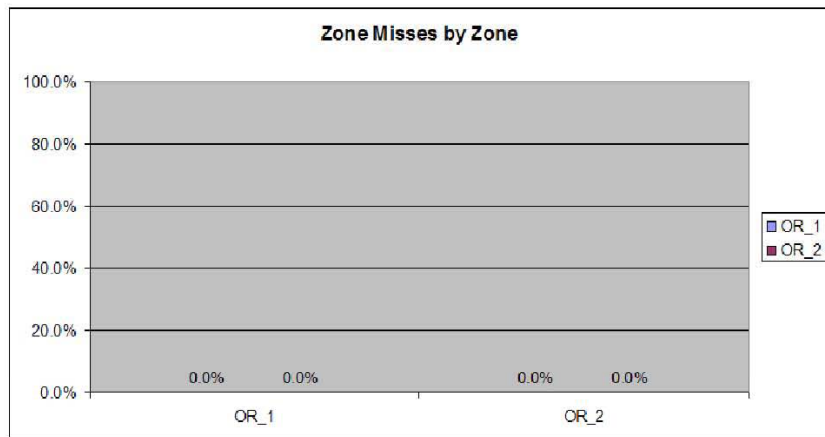


Figure 21: High Volume Tests; Reported Zone vs. Actual, Sonitor

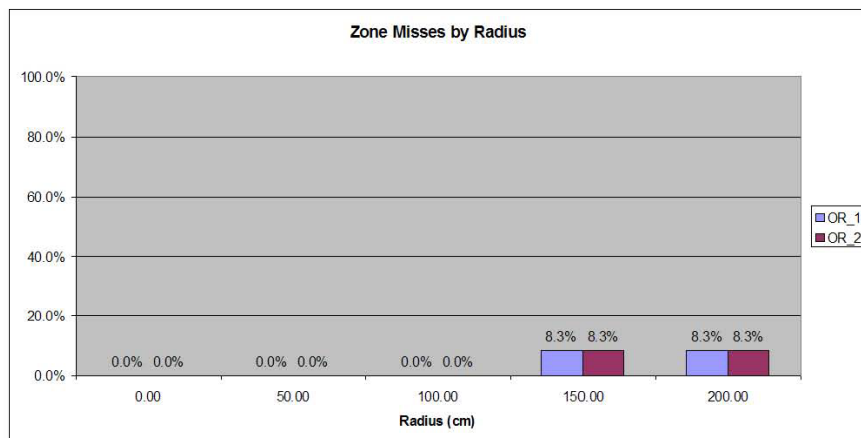


Figure 22: Colocation Tests; Reported Zone vs. Actual, Sonitor

## C Radianse

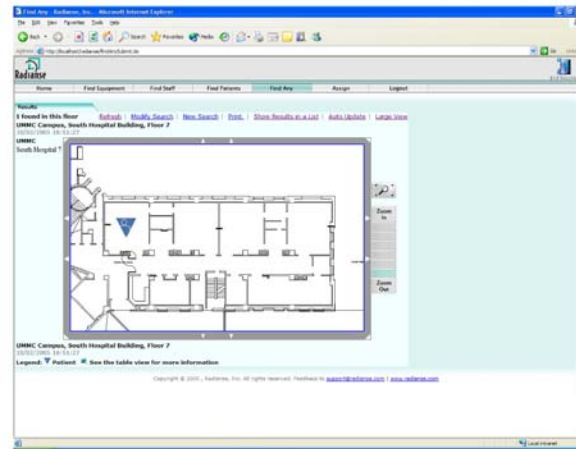


Figure 23: Application Screen, Radianse



Figure 24: Tag, Radianse



Figure 25: Reader, Radianse

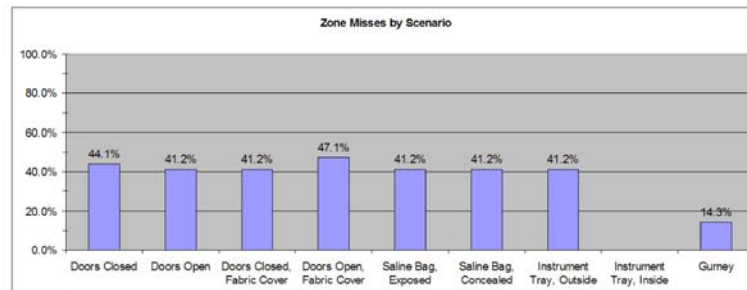


Figure 26: Single Tag Tests; Reported Zone vs. Actual, Radianse

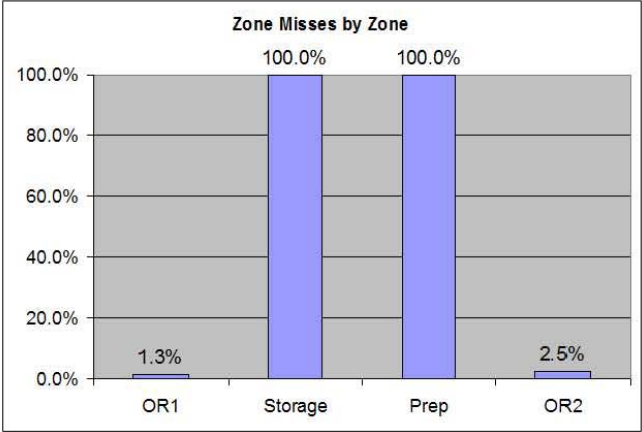


Figure 27: Single Tag Tests; Reported Zone vs. Actual, Radianse

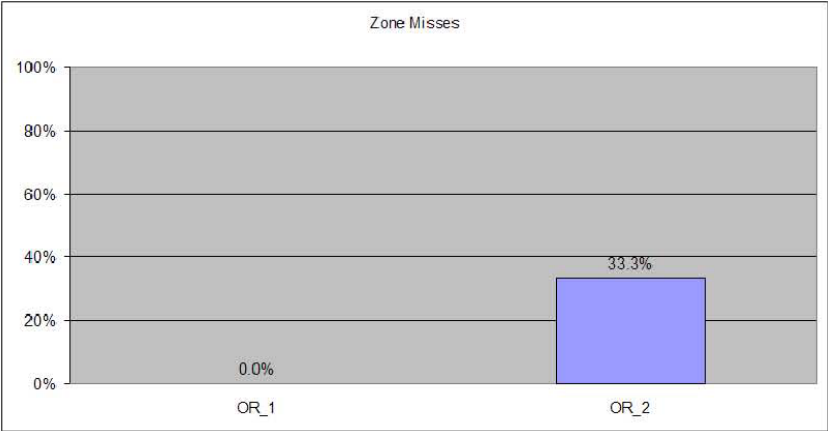


Figure 28: Repeatability Tests; Reported Zone vs. Actual, Radianse

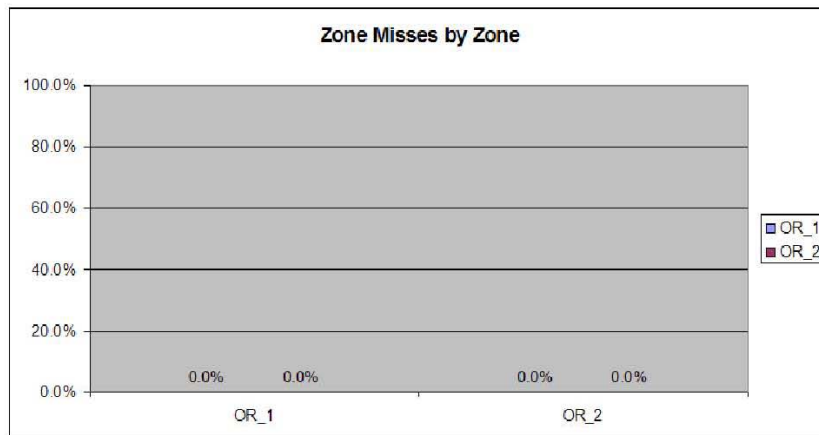


Figure 29: High Volume Tests; Reported Zone vs. Actual, Radianse

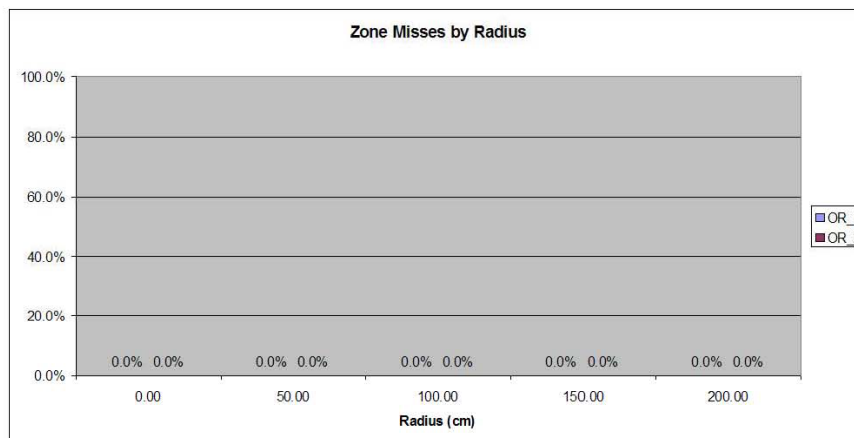


Figure 30: Colocation Tests; Reported Zone vs. Actual, Radianse

## D UbiSense



Figure 31: Tag, UbiSense



Figure 32: Reader, UbiSense

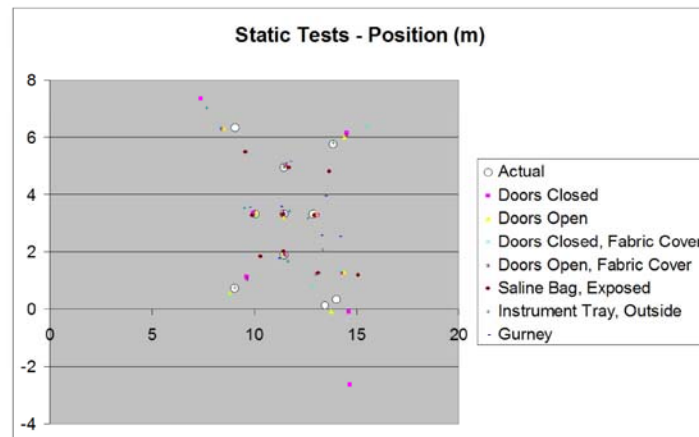


Figure 33: Single Tag Tests; Reported Position vs. Actual, UbiSense

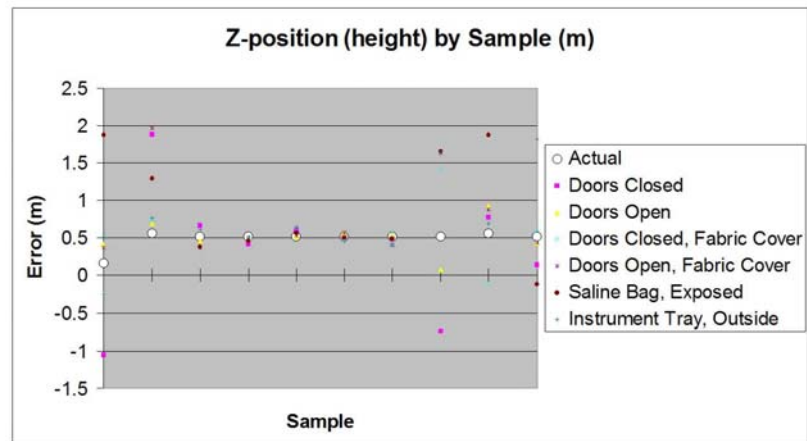


Figure 34: Single Tag Tests; Reported Z-Position vs. Actual, UbiSense

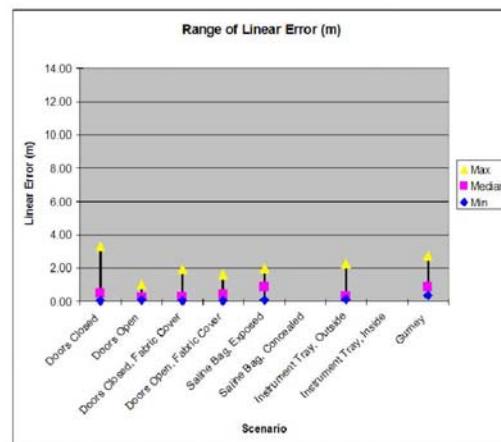


Figure 35: Single Tag Tests; Range of Error in Reported Position, UbiSense



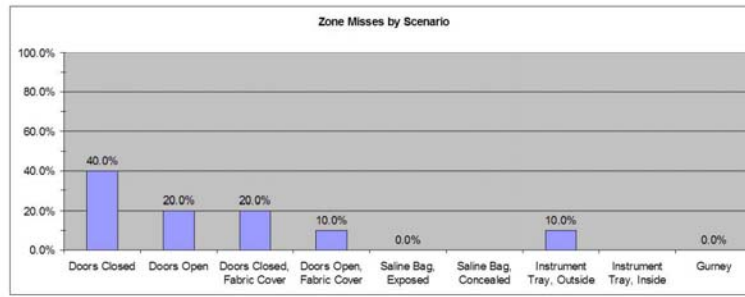


Figure 36: Single Tag Tests; Reported Zone vs. Actual, UbiSense

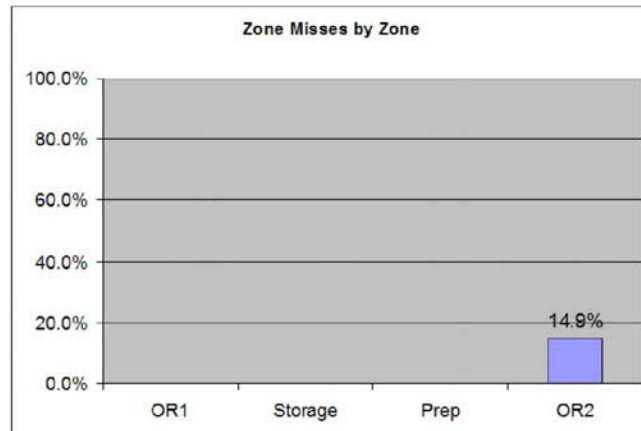


Figure 37: Single Tag Tests; Reported Zone vs. Actual, UbiSense

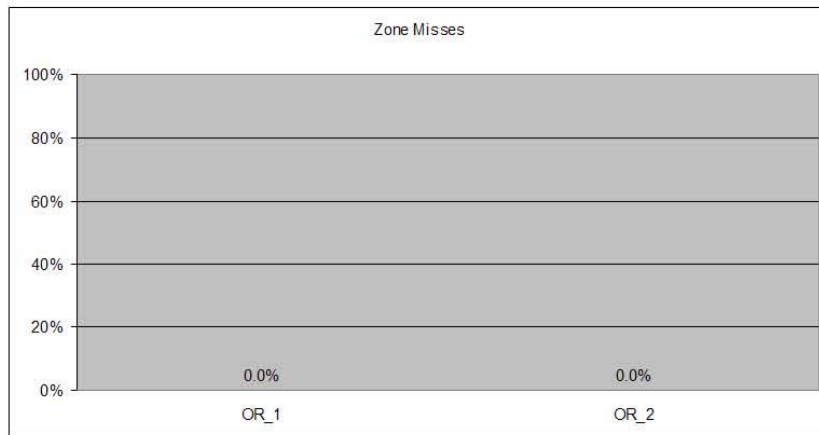


Figure 38: Repeatability Tests; Reported Zone vs. Actual, UbiSense

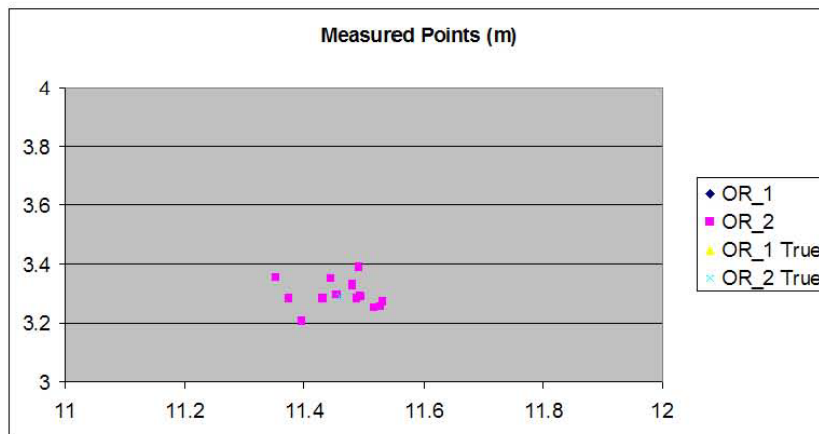


Figure 39: Repeatability Tests; Reported Position vs. Actual, UbiSense

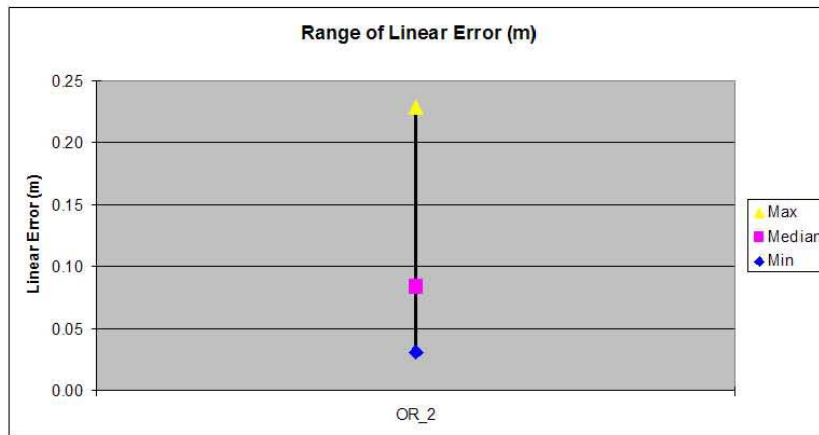


Figure 40: Repeatability Tests; Range of Error in Reported Position, UbiSense

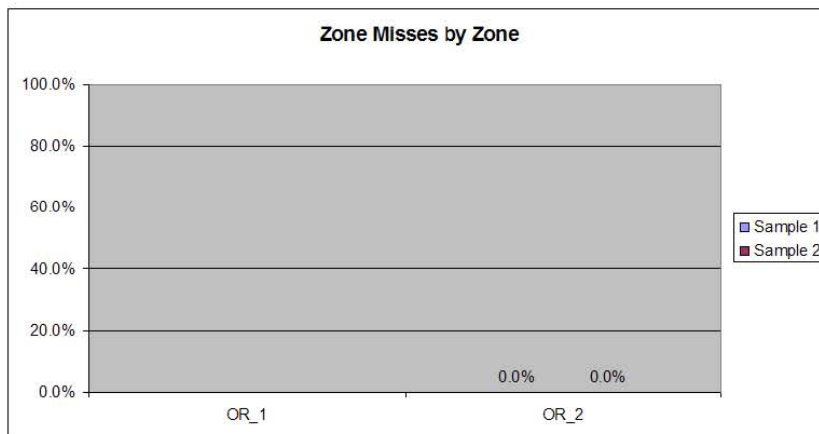


Figure 41: High Volume Tests; Reported Zone vs. Actual, UbiSense

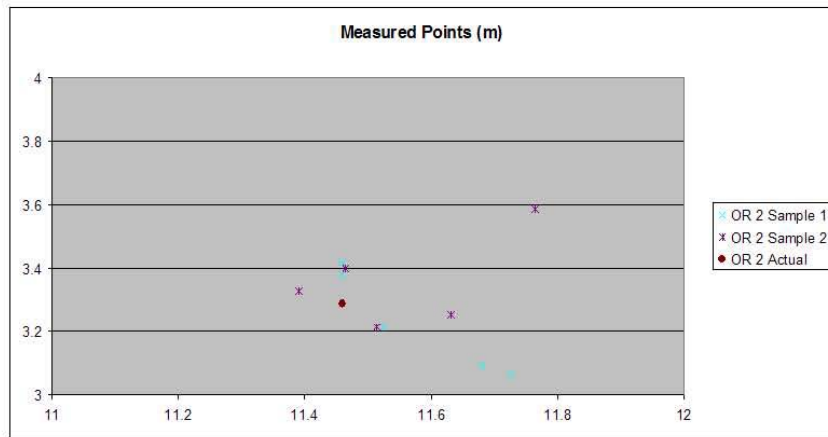


Figure 42: High Volume Tests; Reported Position vs. Actual, UbiSense

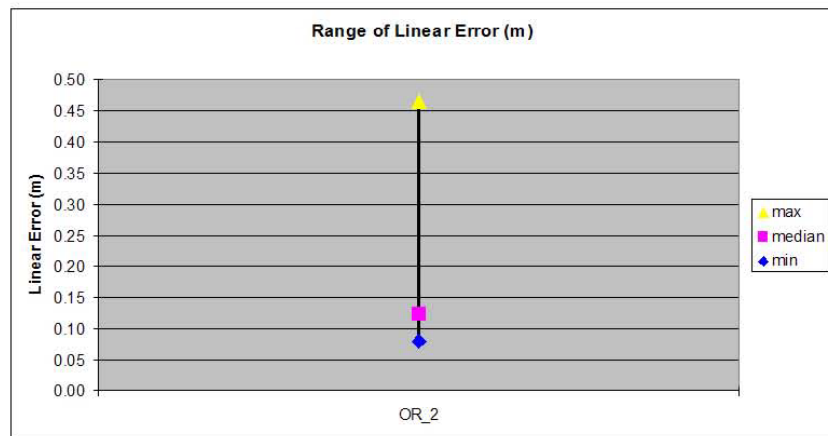


Figure 43: High Volume Tests; Range of Error in Reported Position, UbiSense

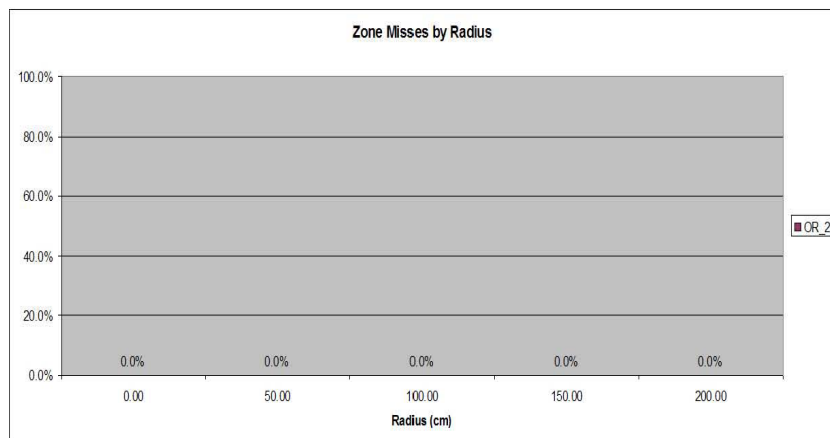


Figure 44: Colocation Tests; Reported Zone vs. Actual, UbiSense

## E Agility Healthcare

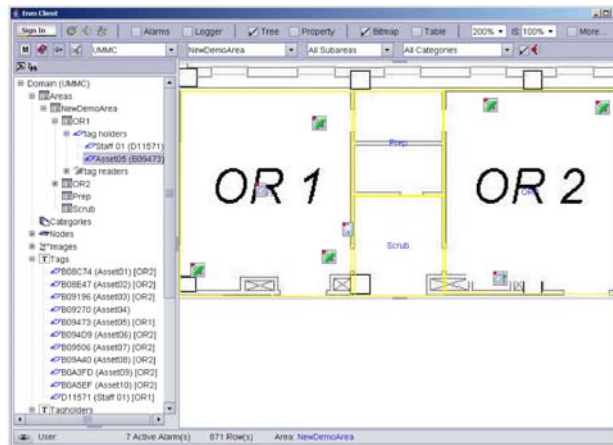


Figure 45: Application Screen, Agility Healthcare



Figure 46: Tag, Agility Healthcare



Figure 47: Zone Reader, Agility Healthcare



Figure 48: Portal Reader, Agility Healthcare

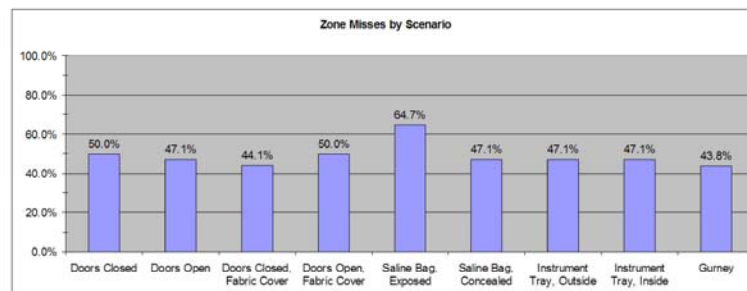


Figure 49: Single Tag Tests; Reported Zone vs. Actual, Agility Healthcare

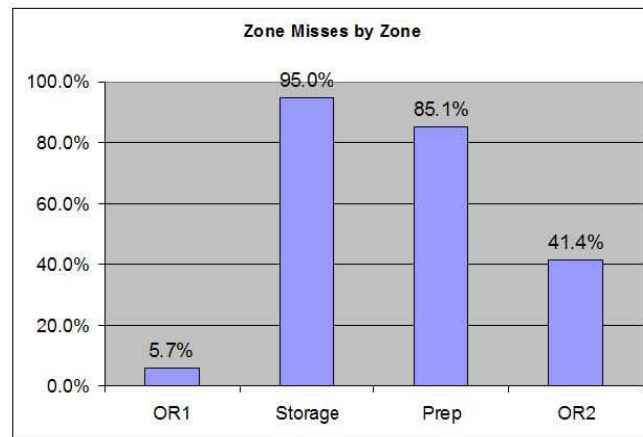


Figure 50: Single Tag Tests; Reported Zone vs. Actual, Agility Healthcare

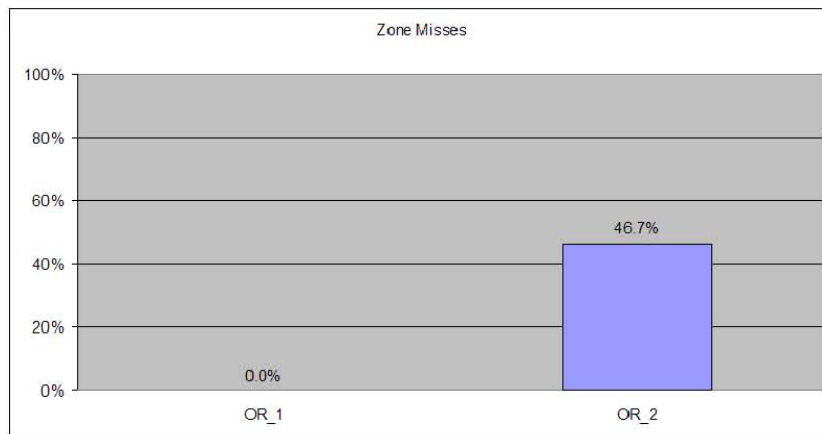


Figure 51: Repeatability Tests; Reported Zone vs. Actual, Agility Healthcare



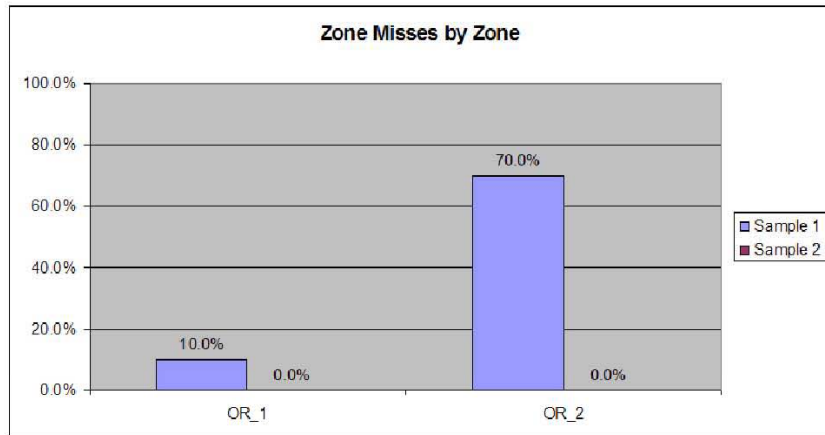


Figure 52: High Volume Tests; Reported Zone vs. Actual, Agility Healthcare

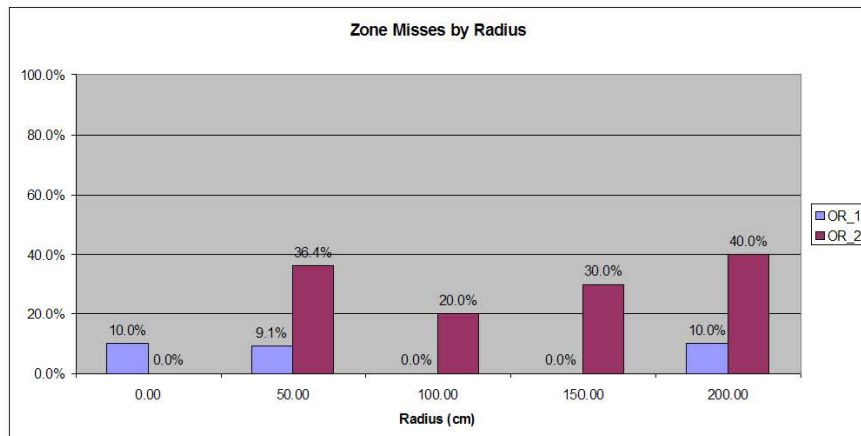


Figure 53: Colocation Tests; Reported Zone vs. Actual, Agility Healthcare

## F Exavera



Figure 54: Staff Tag, Exavera



Figure 55: Asset Tag (Reserach Prototype), Exavera



Figure 56: Asset Tag (Production Prototype), Exavera



Figure 57: Reader (VeraFi Unit), Exavera



Figure 58: Reader (Relay Unit), Exavera

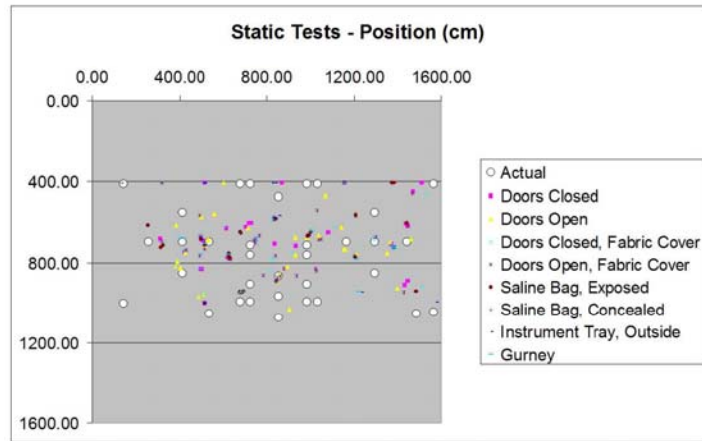


Figure 59: Single Tag Tests; Reported Position vs. Actual, Exavera

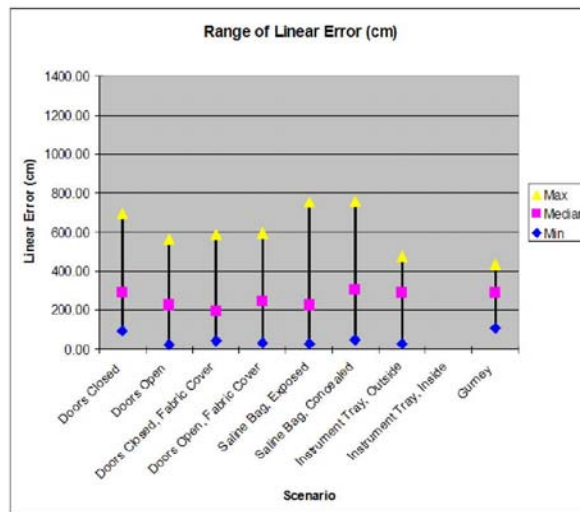


Figure 60: Single Tag Tests; Range of Error in Reported Position, Exavera



Figure 61: Single Tag Tests; Reported Zone vs. Actual, Exavera

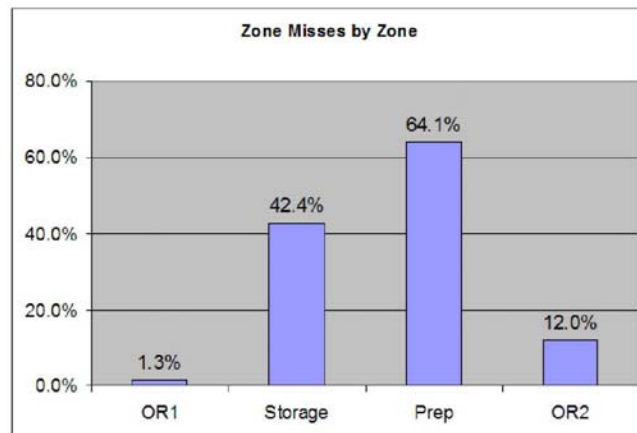


Figure 62: Single Tag Tests; Reported Zone vs. Actual, Exavera



Figure 63: Repeatability Tests; Reported Zone vs. Actual, Exavera

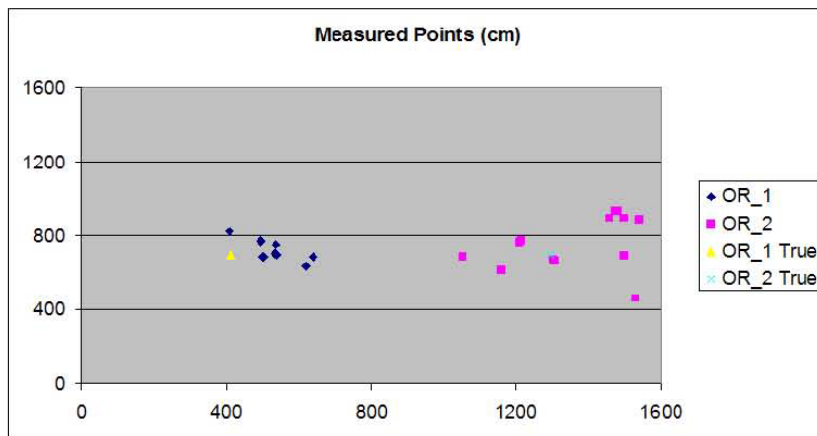


Figure 64: Repeatability Tests; Reported Position vs. Actual, Exavera

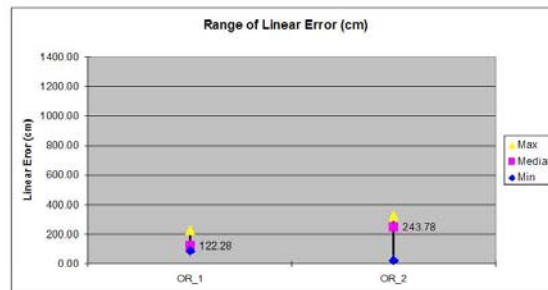


Figure 65: Repeatability Tests; Range of Error in Reported Position, Exavera

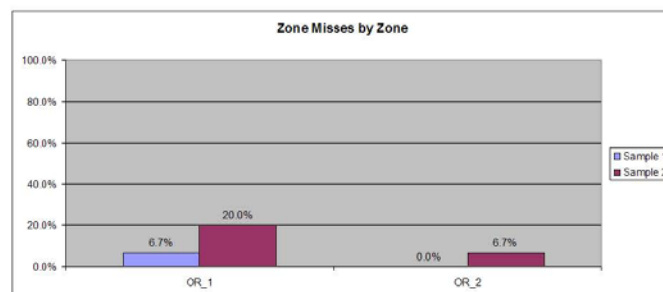


Figure 66: High Volume Tests; Reported Zone vs. Actual, Exavera

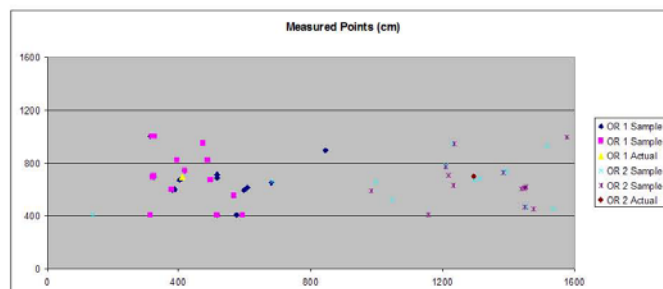


Figure 67: High Volume Tests; Reported Position vs. Actual, Exavera

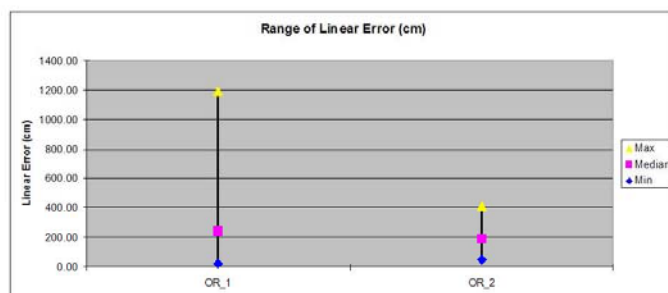


Figure 68: High Volume Tests; Range of Error in Reported Position, Exavera

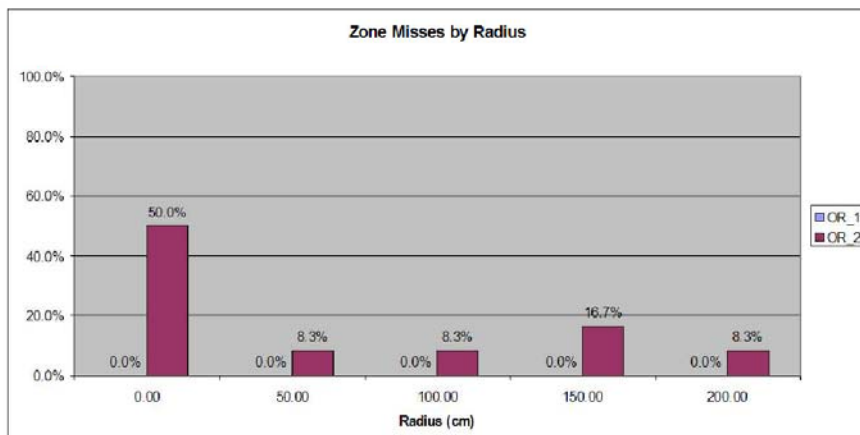


Figure 69: Colocation Tests; Reported Zone vs. Actual, Exavera



## G Test Configurations



Figure 70: Test Scenario S01, Counter-Top Test Point



Figure 71: Test Scenario S03



Figure 72: Test Scenario S05



Figure 73: Test Scenario S06



Figure 74: Test Scenario S07



Figure 75: Test Scenario S08



Figure 76: Test Scenario S09



Figure 77: Test Scenario S11 (Contents of Box)



## **Exhibit 2-A**

# IMAGE REGISTRATION ACCURACY WITH LOW-DOSE CT: HOW LOW CAN WE GO?

*Omkar Dandekar<sup>1,2</sup>, Khan Siddiqui<sup>3</sup>, Vivek Walimbe<sup>4,5</sup>, Raj Shekhar<sup>2</sup>*

<sup>1</sup>Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20740

<sup>2</sup>Department of Diagnostic Radiology, University of Maryland, Baltimore, MD 21201

<sup>3</sup>VA Maryland Health Care System, Baltimore, MD 21201

<sup>4</sup>Biomedical Engineering Center, The Ohio State University, Columbus, OH 43210

<sup>5</sup>Dept of Biomedical Engineering, The Cleveland Clinic Foundation, Cleveland, OH, 44195

## ABSTRACT

Image-guided interventions are known to lead to improved outcomes and significantly faster patient recovery as compared with conventional open, invasive procedures. Common intraoperative imaging techniques such as endoscopy and fluoroscopy are two-dimensional (2D), and provide a 2D representation of the 3D anatomy. Use of recently emerged multislice computed tomography (CT) can facilitate 3D visualization of anatomy during an

intervention. The speed and dimensionality of these CT scanners make their use in image-guided interventions technically feasible. For clinical acceptance, however, the net radiation dose to the patient and the interventionist must be minimized. This article suggests a strategy to reduce radiation dose and describes an evaluation scheme to identify the optimal dose which does not sacrifice the specificity of the image-guided procedure. Our work indicates at least a tenfold reduction in radiation dose.

## 1. INTRODUCTION

laparoscopic cholecystectomy).

Minimal invasiveness is the defining feature and the primary motivator for the adoption of image-guided interventions. Other equally important advantages of image-guided interventions are improved outcomes, shorter procedures and a quick recovery [1]. The current use of image guidance can be found in neurosurgery, focal cancer ablative therapies, and radiosurgeries and radiotherapies. In many clinical applications, image-guided intervention has become the standard of care (e.g.,

To compensate for this misregistration, the anatomy needs re-imaging immediately prior to and during the intervention. Registering preoperative images then with the intraoperatively collected images provides valuable information for updating operative/surgical navigation plan. An interesting aspect of this approach is a possibility to augment the current anatomy, with supplementary information (such as contrast) from the preoperative scan, which will tremendously aid navigation and localization.

Image-guided procedures are usually planned on preoperative CT/magnetic resonance (MR) images. However, due to normal tissue motion, the anatomy at the time of the surgical procedure generally differs from the anatomy at the time of planning. Such misregistration between pre- and intraoperative anatomy renders the planned treatment and navigation, based solely on preoperative data, inaccurate and unsafe.

With the advent of multislice CT (up to 64), intraoperative imaging is now possible with computed tomography. Comparatively, magnetic resonance imaging (MRI) continues to lag in speed, while use of ultrasound is limited due to poor image quality and, in image-guided surgeries specifically, the difficulty of scanning across the pneumoperitoneum with CO<sub>2</sub> insufflation. Radiation exposure to the patient and the interventionist, however, continues to be a concern with using CT, and must be minimized.

Our primary radiation dose reduction strategy is to acquire a standard-dose CT image preoperatively and scan the dynamic operative field subsequently using low-

dose CT. Using high-speed nonrigid 3D image registration techniques [2] our group has developed, the preoperative CT image can be registered to low-dose intraoperative CT images. Registered preoperative CT will show the dynamic intraoperative anatomy and will substitute the low-dose CT images. These diagnostic quality images can be 3D rendered and used for intraoperative guidance and navigation. Capability of viewing hidden structures using CT together with the additional capability of virtually inserting tracked tools in the 3D renderings will add a new dimension to image-guided interventions.

Our proposed dose reduction strategy necessitates the determination of the lowest acceptable dose for

Volume subdivision-based methods with mutual information (MI) as the similarity measure have been shown to be effective for 3D elastic image registration. Normalized mutual information (NMI), in particular, is robust, intensity and overlap independent and hence ideally suited for multimodality image registration. Walimbe et al. [3,4] have reported a volume subdivision-based elastic registration algorithm with NMI as a similarity method. This algorithm has been clinically validated and shown to achieve registration accuracy on the order of the voxel size [5].

This aforementioned algorithm was used to register the standard-dose (200 mAs, 120 KV) CT images with a series of simulated low-dose (150 to 10 mAs, 120 KV) CT images. The accuracy of registration was evaluated as described in the following section.

### 3. EVALUATION OF REGISTRATION ACCURACY

Quantifying the accuracy of elastic registration, in general, is a very difficult task due to the lack of an established standard. It is, however, necessary to judge the registration accuracy in the proposed application to

#### 3.2. Creating anatomically realistic deformations

Human body, and abdominal organs and tissues in particular, undergo elastic deformation during day-to-day activities, respiratory and cardiac cycles, etc. These deformations manifest as misregistration between pre- and intraoperative scans. Further misalignments are introduced due to differences in patient position during imaging as well as different scan parameter settings. In order to create a realistic deformation that incorporates all these effects, it is necessary to estimate this deformation

intraoperative CT, which permits accurate image registration. This article describes an evaluation scheme to judge the registration accuracy with low-dose CT. The following sections describe this method and the results in detail.

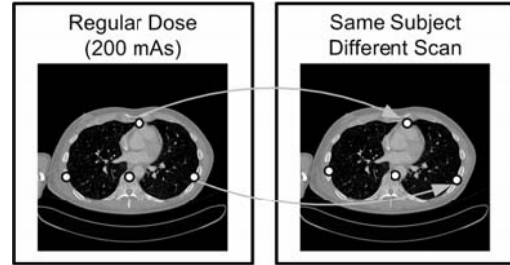


Figure 1 : Generating anatomically realistic deformation

## 2. REGISTRATION ALGORITHM

determine the optimal dose that does not sacrifice the precision of an image-guided procedure. Since our images happened to be registered to begin in the present case, our validation strategy has been to test how well the algorithm recovers a user-introduced, known elastic deformation. Our overall strategy can then be described in three main steps: 1) introduce the same known deformation in low-dose CT images 2) elastically register the (preoperative) standard-dose image with the (intraoperative) low-dose images 3) Compare the transformation field obtained after image registration with the original, user-introduced deformation field to calculate the registration accuracy at various doses.

### 3.1. Generating low dose CT images

Low dose images corresponding to a standard dose abdominal scan were generated using *syngo*-based Somaris/5 simulator from Siemens. This simulator models the noise and attenuation effects at lower radiation doses and can generate low-dose equivalent images from an input standard-dose image. The performance and accuracy of this simulator has been previously reported [6]. This approach ensures that scans at all radiation doses represent exactly the same anatomy.

from scans of the same anatomy taken on different days, thus ensuring sufficient temporal separation.

Deformation vectors between homologous anatomical landmarks in such two scans represent the local misalignment at those landmarks. Using these vectors, deformation field for the entire anatomy was approximated using thin-plate spline (TPS)-based interpolation.

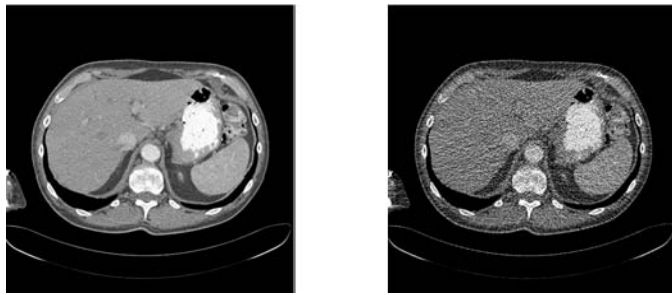
### 3.3 Applying deformation and image registration

The deformation field generated using TPS-based interpolation is applied to the low-dose images. This involves resampling of the low-dose image on to a regular grid using the mapping provided by the deformation field. The registration between the deformed low-dose images

#### 4. MATERIALS AND METHODS

An abdominal scan acquired under clinical settings at the standard dose (200 mAs) was selected for the study. Dose correction feature of the scanner, which automatically adjusts the dose based on the anatomy to be scanned, was turned off for this scan to keep the dose consistent across all the slices. CT image acquired measured  $256 \times 256 \times 300$  voxels, with voxel dimensions of  $1.56 \text{ mm} \times 1.56 \text{ mm} \times 1.5 \text{ mm}$  and field of view restricted mostly to lower thorax and abdomen.

Four different CT scans of the same subject acquired



**a b c Figure 2 : Corresponding slices from (a) 200mAs and (b) 10 mAs scan, (c) 10 mAs scan after anisotropic diffusion filtering**

Treating the standard-dose scan as a reference, we generated eleven low-dose scans (at 10, 15, 20, 25, 30, 40, 50, 70, 85, 100, 150 mAs) using the Somaris/5 simulator. 10mAs was the lowest setting possible for the current version of the simulator. Each of these scans

We registered these deformed and preprocessed low-dose images (reference image) with the original standard-dose image (floating image) using the registration algorithm described earlier. Optimized alignment of the floating and reference image yields a registration field ( $RF_i$ ) which maps each reference image voxel into floating image coordinate space. Comparison of this registration field ( $RF_i$ ) with the original deformation field ( $DF_j$ ) that was introduced is used to judge the registration accuracy for each dose.

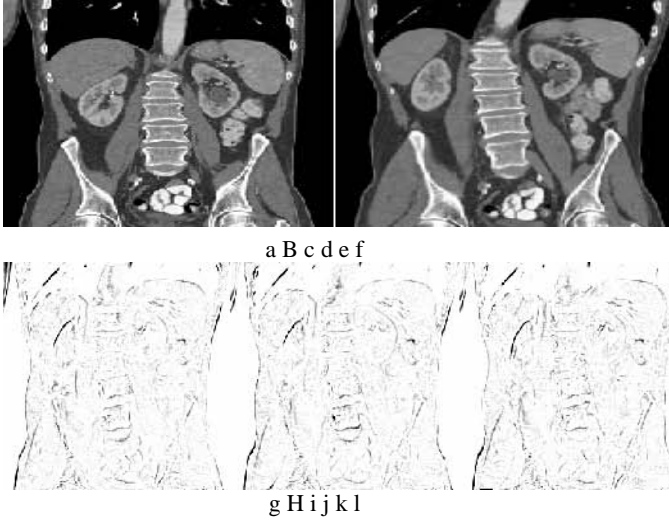
and the original standard-dose image yields a registration field, which reverts the induced deformation. The difference between the registration field and the induced deformation field provides a measure of registration accuracy at that dose.

at different times within a span of one to sixty days were used for creating anatomically realistic deformations. These prior scans had dimensions and resolution of  $256 \times 256 \times 280$ -315 and  $1.48 \text{ mm} \times 1.48 \text{ mm} \times 1.5 \text{ mm}$  respectively. Based on a predetermined, well-described list of 32 anatomical landmarks, a clinical expert identified and marked a set of homologous points in all the CT scans (one standard-dose scan and four older scans). Based on the expert-defined landmarks, thin-plate spline (TPS)-based starting deformation fields ( $DF_1$ ,  $DF_2$ ,  $DF_3$ , and  $DF_4$ ) corresponding to each prior scans were generated.

(including the standard dose scan) was deformed using the registered deformation fields ( $DF_1$ ,  $DF_2$ ,  $DF_3$ , and  $DF_4$ ) described earlier.

The low-dose scans show the exact same anatomy as in the standard dose scan but are characterized by a high quantum noise. These scans may not be acceptable for diagnostic purposes, but they contain sufficient information regarding the current anatomical state. Corresponding standard-dose and a low-dose scan at 10 mAs are shown in Fig 2 a-b. Using low-dose images, as is, for image registration will cause the dispersion of the mutual histogram (due to noise, in an otherwise uniform structure) leading to poor image registration. Anisotropic diffusion filtering has been shown to be an effective processing step prior to advanced image processing [7]. Figure 2c shows the improvement in the visual quality of the low-dose scans after preprocessing.





**Figure 3 : Registration errors with low-dose CT; (a) original image, (b) image after known realistic deformation, (c) corresponding difference image, (d-i) difference**

### 5.1. Qualitative Evaluation

Result of the qualitative evaluation using one of the deformations ( $DF_1$ ) is shown in Fig. 3. Visually correct registration of the standard-dose image with the deformed images at various low doses (evident from the reduced features in the difference image) demonstrates the feasibility of elastic registration at low CT doses. Inter-registration errors indicate that registration results at lower doses are comparable to those obtained using a standard dose.

### 5.2. Quantitative Evaluation

The process of elastic registration attempts to recover any misalignment between the reference and floating images. A perfect registration will completely recover this misalignment and yield an elastic transformation field that is identical to the deformation field representing the original misalignment. A comparison between these two fields can be used as a performance index of the registration algorithm.

images after registration with low-dose CT (for doses 200, 100, 50, 30, 20, 10 mAs, respectively) with respect to the deformed image, (j-l) registration errors between images registered at scan doses of 50, 30, 10 mAs with respect to that at standard dose (200 mAs).

**Table 1 : Average registration error (mm) with various CT doses**

	200	100	50	30	20	10	5	2
Deformation 1	2.12	2.09	2.07	2.06	2.04	2.05	2.02	2.01
Deformation 2	2.22	2.21	2.18	2.17	2.16	2.15	2.14	2.13
Deformation 3	2.19	2.18	2.14	2.11	2.08	2.06	2.03	2.02
Deformation 4	2.09	2.09	2.08	2.07	2.05	2.04	2.02	2.01

Radiation Dose (mAs)

10 15 20 25 30 40 50 70 85 100 150 200

## 5. RESULTS AND DISCUSSION

For this application, the deformation field introduced ( $DF_i$ ) is known at every voxel. The volume subdivision-based elastic registration algorithm generates the transformation field ( $RF_j$ ), which provides the transformation at every voxel in scan with dose  $j$ . The average of the magnitude of the vector differences between these two fields is reported in Table 1. This average was calculated over the region of the image which contains sufficient part of the subject and hence information to yield meaningful registration. The regions of the image which contain no information (very low entropy) (e.g. black areas surrounding the subject) are masked out using a simple threshold operation.

The results show a maximum error of 11% and 9% at the doses of 10 mAs and 20 mAs, respectively. The minimum errors at these doses are 6% and 5%, respectively. As expected, the average error improves steadily with dose. Primary causes of the baseline registration error are the resolution of the images and lowest subvolume size of the registration algorithm. Our future work will involve evaluating registration accuracy using image with finer resolution and potentially smaller subvolume sizes.

## 6. CONCLUSIONS

We have demonstrated successful registration of standard-dose abdominal CT images with lower-dose images of the same anatomy. Even at 10 mAs, the smallest dose, the registration accuracy achieved was comparable to that achieved at the standard dose. Our results demonstrate ten- to twenty fold reduction in

radiation dose with the use of low-dose CT. Reduction of radiation dose to safe levels is highly significant in that it enables navigating interventions using more powerful, multislice CT. Our work promises to bring in a new level of sophistication and accuracy in image-guided interventions through the introduction of true 3D visualization possible through safe, volumetric CT imaging of the intraoperative anatomy.

## 7. REFERENCES

- [1] Harrell AG, and Heniford BT, "Minimally invasive abdominal surgery: lux et veritas past, present, and future," American Journal of Surgery, 190 (2): 239-243 Aug 2005.
- [2] Castro-Pareja CR, Shekhar R, "Hardware acceleration of mutual information based 3D image registration," Journal of Imaging Science and Technology 2005;49:105-113.
- [3] Walimbe V, Shekhar R, "Automatic elastic image registration by interpolation of 3D rotations and translations from discrete rigid-body transformations," Medical Image Analysis (accepted for publication).
- [4] Walimbe V, Zagrodsky V, Raja S, Bybel B, Kanvinde M, and Shekhar R, "Elastic registration of three-dimensional whole body CT and PET images by quaternion based interpolation of multiple piecewise linear rigid-body registrations," Proc SPIE 2004, 5370:1191-1228.
- [5] Shekhar R, Walimbe V, Raja S, Zagrodsky V, Kanvinde M, Wu G, and Bybel B, "Automated 3-Dimensional Elastic Registration of Whole-Body PET and CT from Separate or Combined Scanners," J Nucl Med 2005, 46:1488-1496.
- [6] Sennst D, Kachelriess M, Leidecker C, Schmidt B, Watzke O, and Kalender W, "An extensible software-based platform for reconstruction and evaluation of CT images," RADIOGRAPHICS, 24 (2): 601-613 Mar 2004.
- [7] A. Dorati, C. Lamberti, A. Sarti, P. Baraldi, and R. Pini, "Preprocessing for 3D echocardiography," Computers in Cardiology 1995: IGEA, Modena, Italy, 565-568.



## Development of Continuous CT-Guided Minimally-Invasive Surgery

*Raj Shekhar, Omkar Dandekar, Matthew Weiner, Ivan George, Patrick Malloy, Reuben Mezrich, Adrian Park*  
Department of Diagnostic Radiology, University of Maryland School of Medicine

## I. Problem

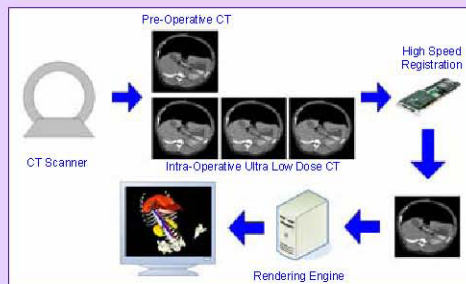
Laparoscopes used to visualize internal anatomy and navigate minimally-invasive surgeries have limited visualization capability. They offer a restricted field of view, relatively flat representation of 3D anatomy and poor vessel contrast. More important, they can display only the superficial surfaces. A surgeon is unable to see **around** or **inside** a structure, limiting the precision of current-generation laparoscopic surgeries. An improved awareness of the 3D operative field is a long-standing need of laparoscopic surgeons that laparoscopes are fundamentally limited in meeting.

## II. Solution

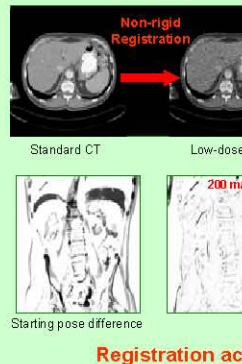
Continuous multi-slice CT of the operative field can be used as a supplementary tool to navigate laparoscopic surgeries. CT can offer large field of view and additional information such as contrast-enhanced vessels. Unlike laparoscopes, CT can help visualize hidden structures. For routine clinical use, however, the net radiation dose must be reduced.

## III. Dose Reduction Strategy

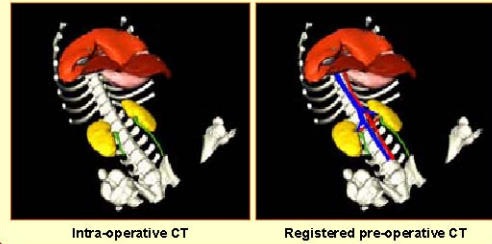
Use ultra low-dose intra-operative CT to modify (through non-rigid registration) contrast-enhanced, standard preoperative CT



## IV. Non-rigid Registration



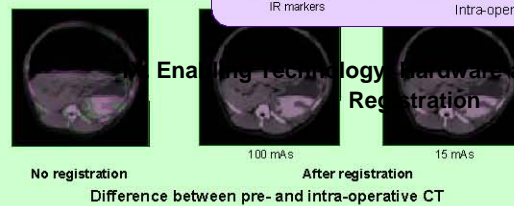
## VII. Intra-operative Vessel Enhancement



## VIII. Tool Tracking

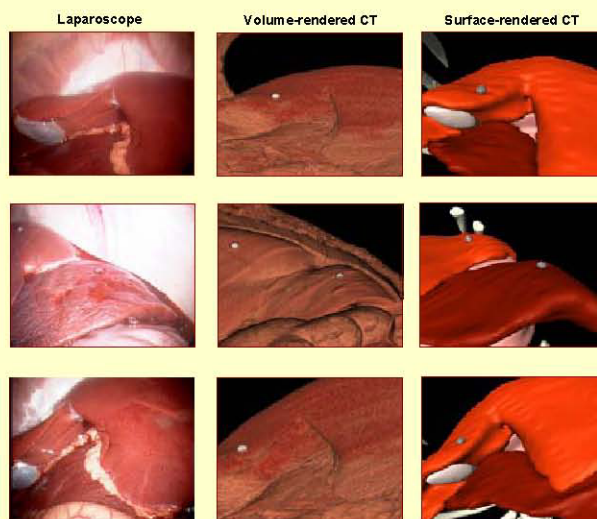


## V. Non-rigid Registration



## Enabling Technology: Hardware-accelerated Image Registration

## VI. Comparison of intra-operative visualization





Implementation	Rigid	Nonrigid
Software	30-40 min	2-12 hours
1 <sup>st</sup> Gen Hardware (30X)	60-80 sec	-
2 <sup>nd</sup> Gen Hardware (100X)	18-24 sec	1.2-7.2 min

### X. Conclusion

We introduce a new volumetric navigation paradigm for minimally-invasive surgeries using continuous multi-slice CT. CT enables visualization of internal structures and blood vessels intra-operatively. Further, >10-fold reduction in radiation dose is possible with pre- to intra-operative CT image registration, which can be accelerated for real-time surgical guidance. Our research has the potential to help improve the precision of laparoscopic surgeries, reduce complications, and expand the scope of minimally invasive surgeries to beyond its current 15% share of all surgeries.

### XI. References

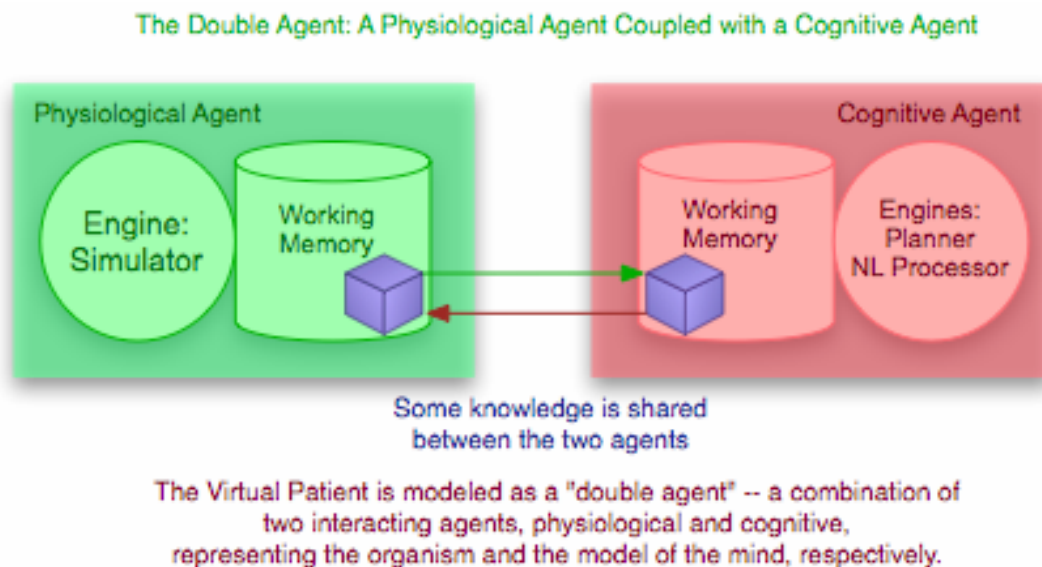
1. Dandekar O, Siddiqui K, Walimbe V, Shekhar R, "Image registration accuracy with low-dose CT: How low can we go?", IEEE ISBI 2006 Accepted
2. Castro-Pareja CR, Shekhar R, "Hardware acceleration of mutual information-based 3D image registration", Journal of Imaging Science and Technology 49(2):105-113, 2005

## Appendix 3-A

### **Double Agent: An Environment for Automatic Tutoring of Medical Students Using Simulation Involving a Combination of a Physiological Software Agent, a Cognitive Software Agent, Software Agents Representing Members of the Medical Team and a Combination of Human and Software Agents for Mentoring.**

Sergei Nirenburg, Marjorie McShane, Stephen Beale and Thomas O'Hara (UMBC)  
Bruce Jarrell, George Fantry and John Raczek (UMB)

This is a very brief description of the overall concept.

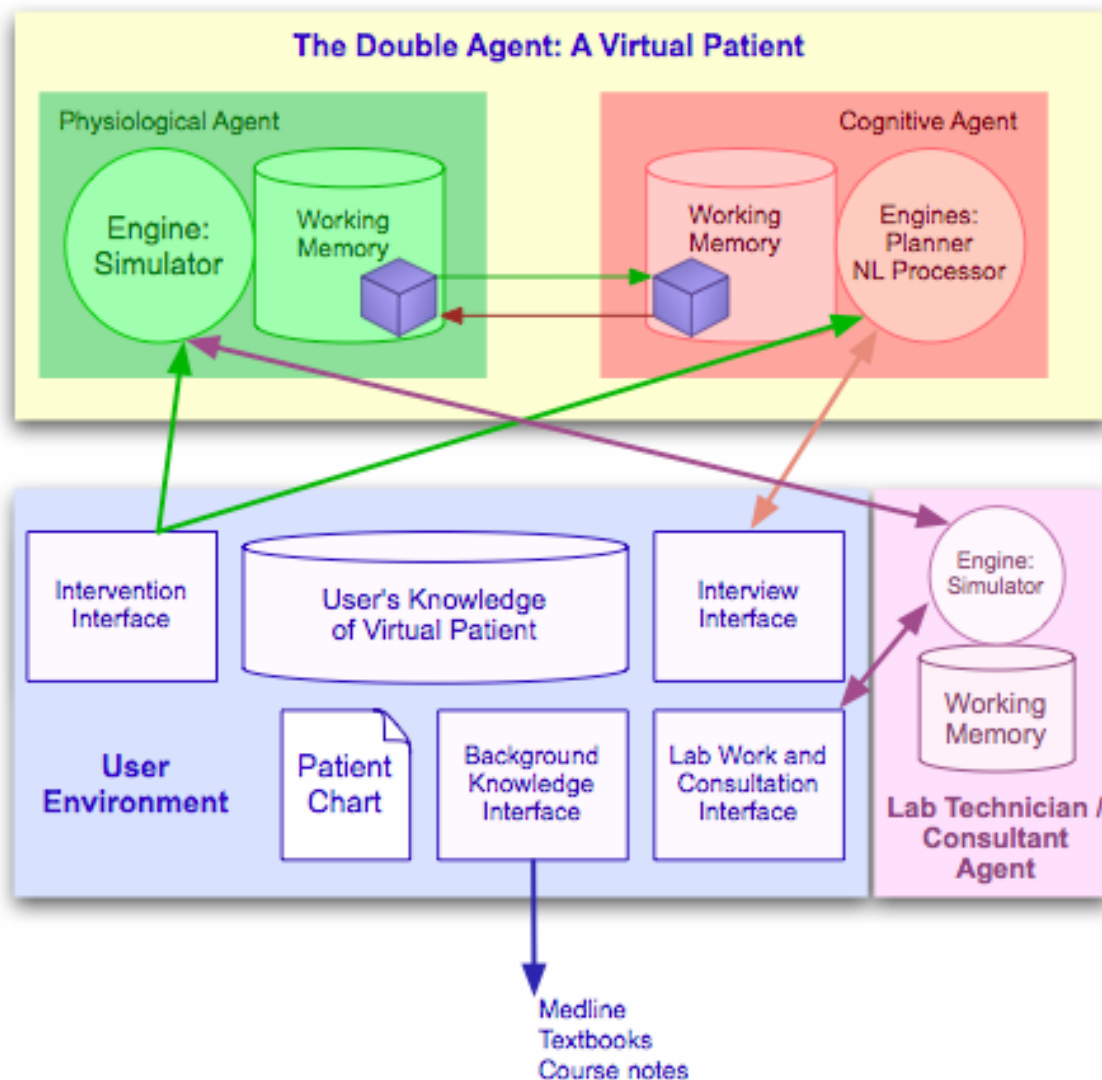


The virtual patient is realized as a “double agent,” a combination of a physiological and a cognitive software agent. The specific nature of the agents is determined by a specific set of values of descriptive properties defined in the underlying world knowledge base, the ontology. The property values specific to a particular virtual patient are represented in the working memory of the physiological and the cognitive agent. The working memory of the physiological agent contains properties that are typically not consciously manipulated by people. The working memory of the cognitive agent contains the properties that are typically consciously manipulated by people as well as memory of past actions, conversations and other events and states. A subset of Double Agent’s property values can be accessed and manipulated by both the physiological and the cognitive agent.

The physiological agent operates through a simulator that modifies the values of properties accessible to the agent as a result of applying the causally and temporally organized knowledge about physiological processes – normal, pathological and those induced by medical intervention. The cognitive agent operates by communicating with the user (either through natural language or through a menu-driven interface), by triggering conscious actions in response to outside requests or internal planning resulting from changes in its physical and mental states and by recording the



events that occurred and the states in which the agent found itself for future use in planning and communication.



**The user can interview the cognitive agent, learn about the state of the virtual patient by ordering lab work and diagnostic consultations and intervene with the physiological agent by prescribing treatment**

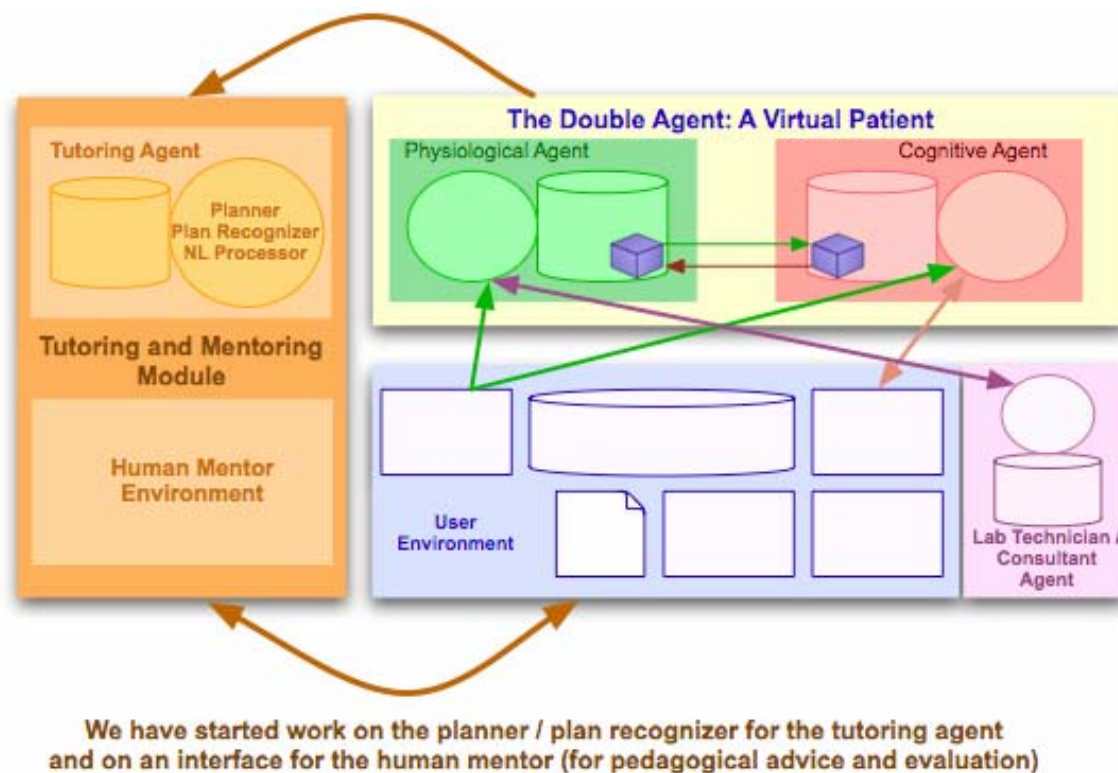
The Double Agent is used by a human user, the medical student. The latter operates in an environment that consists of communication interfaces:

- an intervention interface for communicating events to the physiological agent (e.g., “take this pill with a glass of water”) and to the cognitive agent (e.g., “avoid eating chocolate or fatty foods”);
- an interview interface that supports the simulation of an interview with the patient during an office visit;

- a lab work and consultation interface through which the user orders various lab tests and solicits opinions of specialists; and
- a background knowledge interface that allows the user to access reference materials.

As a result of the patient interview, the lab work and the specialists' opinions, the student obtains knowledge about a subset of property values of the Double Agent. These values are recorded in the User's Knowledge of Virtual Patient repository. The patient's chart is filled out partially from those values and partially by the user who can add arbitrary notes and comments.

The lab technician / specialist agent obtains requests from the user, queries the physiological agent, obtains the necessary property values, adds a conclusion in English about the results and communicates it back to the user, together with the numerical results.



A human mentor can carry out arbitrary communicative actions (e.g., provide hints to the user at any time in the process). The software tutoring agent traces the actions of the user and judges how well they match a set of acceptable sequences of actions for the particular case (a specific virtual patient). On the basis of this judgment, it can communicate advice to the user, through canned messages or by dynamically generating appropriate text and engaging the student in a dialog.

## **Appendix 3-B**





# **The Double Agent**

**Physiological and Cognitive Agents  
for the Support of Developing the  
Cognitive Skills of Medical Students**



## A team of computer scientists and SMEs

**Sergei Nirenburg,  
Marjorie McShane,  
Stephen Beale  
Thomas O'Hara**

Institute for Language and  
Information Technologies  
University of Maryland Baltimore County

**Bruce Jarrell,  
George Fantry,  
David Mallott,  
John Raczek**

University of Maryland  
School of Medicine

The vision:

“Here is a virtual patient. Try to diagnose and treat the problem it has. See how the patient responds. See if another treatment would have been better.”

NOT

“Try to perform this procedure”

NOT EXACTLY

“Let me teach you to diagnose and treat patients”

## The Long-Term Goal of this research is threefold:

1. Creating a **conceptual** computer simulation of a human organism that is also capable of a cognitive function (“the double agent”)
2. Encoding **biological and clinical knowledge** about normal function, disease states, diagnostic tests and treatments in an ontological knowledge base
3. Combining the above simulation capabilities and knowledge in a variety of **applications**, including:



rehearsal of patient care

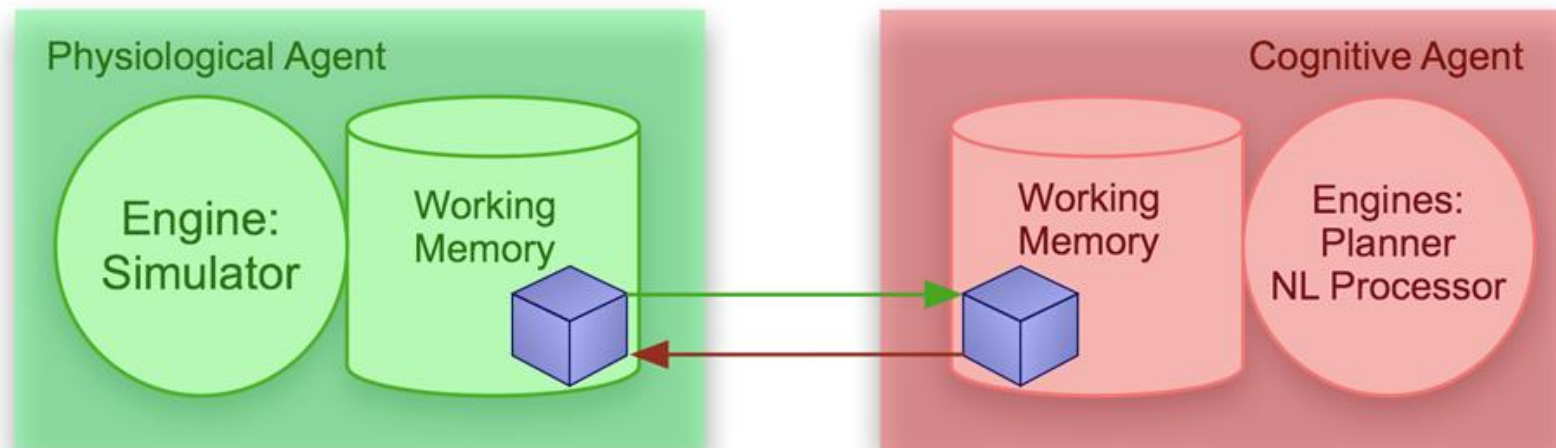


tutoring

## Plan of Talk

1. What is the Double Agent System?
2. How it works: a demo
3. Knowledge sources and knowledge elicitation
4. Discussion

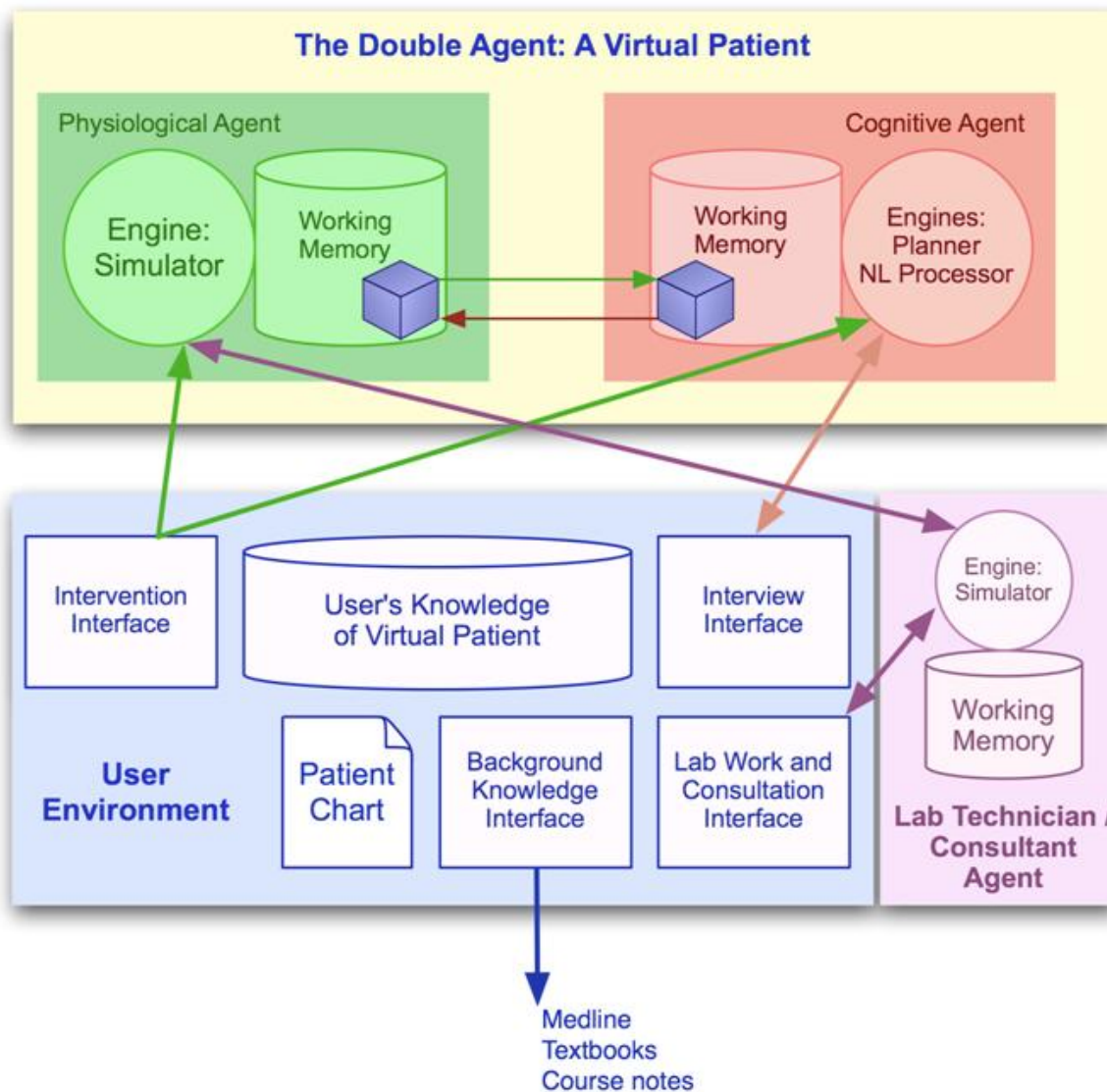
## The Double Agent: A Physiological Agent Coupled with a Cognitive Agent



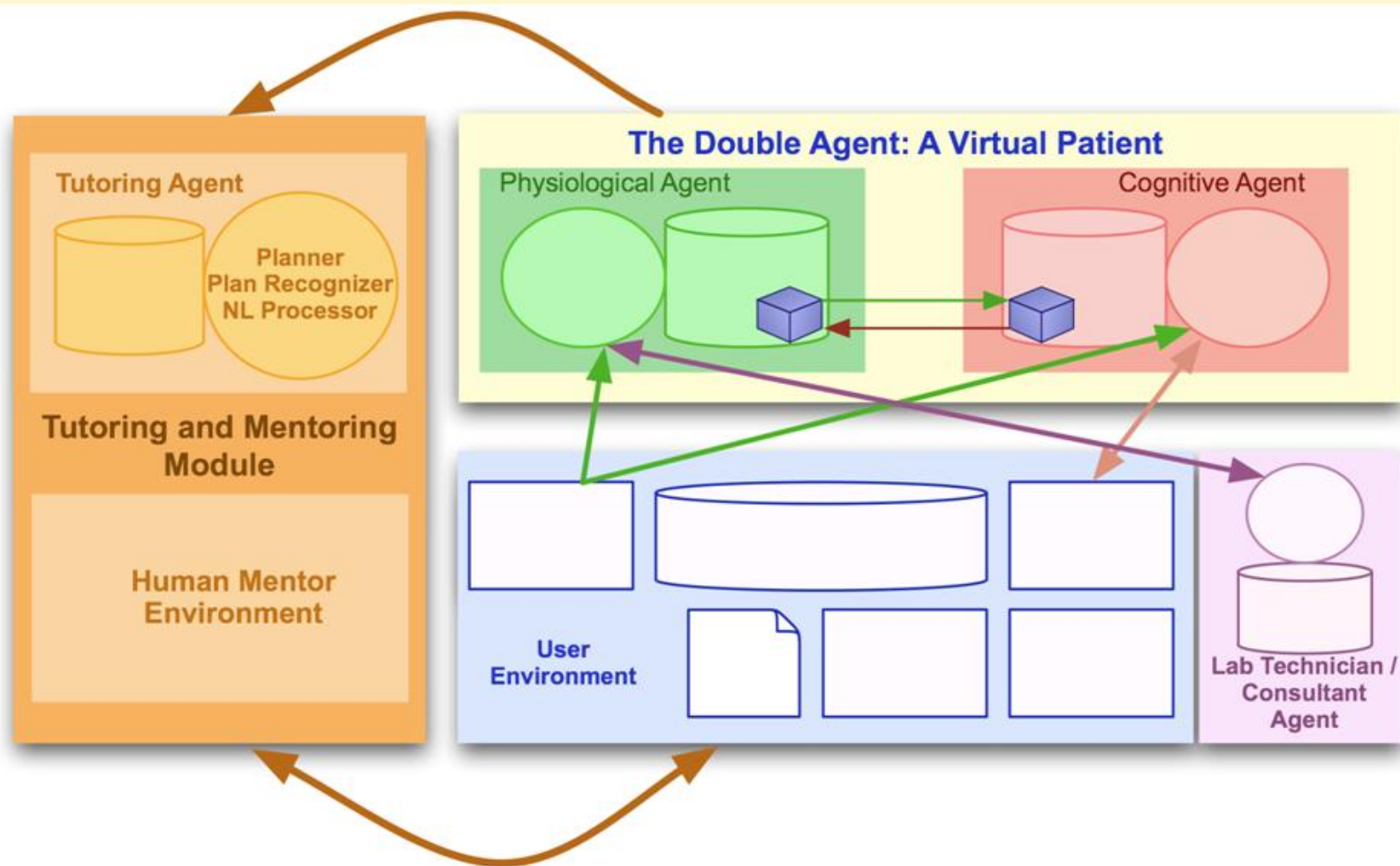
Some knowledge is shared  
between the two agents

The Virtual Patient is modeled as a "double agent" -- a combination of two interacting agents, physiological and cognitive, representing the organism and the model of the mind, respectively.





**The user can interview the cognitive agent, learn about the state of the virtual patient by ordering lab work and diagnostic consultations and intervene with the physiological agent by prescribing treatment**



We have started work on the planner / plan recognizer for the tutoring agent and on an interface for the human mentor (for pedagogical advice and evaluation)



## Plan of Talk

1. What is the Double Agent System?
2. How it works: a demo
3. Knowledge sources and knowledge elicitation
4. Discussion

The Double Agent is not yet a complete tutor.

It is still a set of enabling components:

- A physiological simulator
- A cognitive simulator
- A set of NLP tools (not yet integrated with the simulators)

and the knowledge resources to support the above.

In the immediate future

The work on **knowledge acquisition** will continue: at present, Double Agent deals only with the diseases of the esophagus.

We need to pay much more attention to **interfaces and visualization**.

We also need to **integrate** the tutor (including the NLP capabilities) and conduct actual studies with medical students.

We will be happy to **collaborate** with other groups on visualization issues and on psychology-intensive tasks.

OK, let's now try to run the Double Agent

Note: in the current version, the interface  
and visualization are just minimum-utility  
placeholders

## Plan of Talk

1. What is the Double Agent System?
2. How it works: a demo
3. Knowledge sources and knowledge elicitation
4. Discussion

# Disease Description → Modeling

- Physicians' initial description of GERD (gastroesophageal reflux disease):

“Frequent reflux causes inflammation of the lining of the esophagus that in some patients can lead to erosions, ulcers, esophageal stricture or Barrett’s esophagus. A small percentage of patients with Barrett’s esophagus develop a cancerous tumor.”

- But this doesn’t yet describe a *disease*. Does everyone with reflux have GERD?



# Digging deeper...

When you start probing the physician SMEs, you find that:

Actually, to have GERD, a patient's lower esophageal sphincter (LES) must either have a low basic pressure or be subject to transient relaxations, both of which permit stomach contents to reflux frequently.

If a patient has GERD, the disease can progress along any one of six different paths.

# The 6 disease paths of GERD

Which disease path a patient experiences depends on individual predispositions (e.g., some patients never progress past inflammation).

Inflammation

Inflammation – erosion

Inflammation – erosion – ulcer

Inflammation – erosion – ulcer – peptic stricture

Inflammation – Barrett's esophagus

Inflammation – Barrett's esophagus – tumor



Apart from the basic disease path, patients differ with respect to:

- what environmental factors exacerbate symptoms (e.g., drinking coffee, eating fatty foods, etc.)
- for multi-stage disease paths, how long each stage lasts
- which symptoms the patient experiences, with what intensity and frequency
- how the patient responds to different interventions, from lifestyle modifications to drugs
- and so on...

# GERD: Disease Model

Physiological Property Values	t1 "inflam."	t2 "erosion"	t3 "ulcer"	t4 "stricture"
Total time in reflux	1.2 – 1.44 (1.44)	1.68 – 2.4 (1.92)	2.4 – 3.6 (2.88)	1.68 – 3.6 (3.6)
DeMeester score	10 – 18 (18)	25 – 40 (32)	40 – 60 (48)	25 – 60 (60)
Total time in reflux; no bad habits	.96 – 1.44 (.96)	1.68 – 2.4 (1.92)	2.4 – 3.6 (2.88)	1.68 – 3.6 (3.6)
DeMeester; no bad habits	7 – 18 (7)	25 – 40 (32)	40 – 60 (48)	25 – 60 (60)
Length of erosion (cm.); [end value]		.5 – 4 (2)		
Diam. of erosion (mm.); [end value]		.5 – 3 (1)		
Number of erosions; [end value]		1 – 4 (2)		
Depth of ulcer (mm.); [end value]			1 – 3 (2)	
Diameter of ulcer (mm.); [end value]			4-10 (5)	
Number of ulcers; [end value]			1 – 4 (2)	
Diam. of T10 lumen (cm.); [end value]				.5

Symptoms	t1 "inflam."	t2 "erosion"	t3 "ulcer"	t4 "stricture"
Heartburn frequency (#/day)	3 – 5 (4)	6 – 8 (7)	9 – 10 (9)	6 – 10 (7)
Heartburn severity	.3 – .5 (.4)	.6 – .8 (.7)	.9 – 1.0 (.9)	.6 – 1.0 (.7)
Symptom correlation (for pH monitoring)	0 – 1	0 – 1	0 – 1	0 – 1
Regurgitation freq. (#/day)	3 – 5 (4)	6 – 8 (7)	9 – 10 (9)	6 – 10 (7)

Treatments (does it work? y / n)	t1 "inflam."	t2 "erosion"	t3 "ulcer"	t4 "stricture"
PPI	y / n	y / n	y / n	y / n
H2 Blocker	y / n	y / n	y / n	y / n
Lifestyle Modifications	y / n	y / n	y / n	y / n
LES diameter after TTS dilation				1 - 3

We are now half way to actual knowledge encoding

# Model to Patient Instance

Name: Douglas Jeeves

Basic LES pressure: 10

GERD-irritating substances: chocolate, caffeine

White male; 60 years old

Transient relaxations: yes

t1 – t4 durations (mos.): 18, 10, 8, 14

Physiological Property Values	t1 "inflamm."	t2 "erosion"	t3 "ulcer"	t4 "stricture"
Total time in reflux	1.44	2	2.5	3.5
DeMeester score	18	27	44	50
Total time in reflux; no bad habits	.96	1.9	2.5	3.5
DeMeester; no bad habits	7	26	44	50
Length of erosion (cm.); [end value]		.7		
Diam. of erosion (mm.); [end value]		2		
Number of erosions; [end value]		1		
Depth of ulcer (mm.); [end value]			1	
Diameter of ulcer (mm.); [end value]			4	
Number of ulcers; [end value]			1	
Diam. of T10 lumen (cm.); [end value]				.5

Symptoms	t1 "inflamm."	t2 "erosion"	t3 "ulcer"	t4 "stricture"
Heartburn frequency (#/day)	3	6	9	8
Heartburn severity	.3	.7	.9	.6
Symptom correlation (for ph monitoring)	.7	.77	.6	.5
Regurgitation freq. (#/day)	3	8	8	8

Treatments (does it work? y / n)	t1 "inflamm."	t2 "erosion"	t3 "ulcer"	t4 "stricture"
<b>PPI</b>	y	y	y	y
<b>H2 Blocker</b>	y	y	n	n
<b>Lifestyle Modifications</b>	n	n	n	n
<b>LES diameter after TTS dilation</b>				1.3



# Patient Instance to Script

All GERD patients begin as experiencers of an inflammation event.

The length of the inflammation stage is recorded as  $t_1$ , which represents a time period in simulation.

The physiological property values and symptoms are calculated for each clock cycle from the start value to the end value of each stage.

At the end of each disease stage, the simulator checks the patient's recorded predispositions and launches the subscript for the next stage.

For GERD, the stage after inflammation can either be erosion or Barrett's esophagus. A given patient can have a predisposition only to one of these (or to neither, in which case the inflammation stage continues indefinitely).

## ...to script (cont.)

- If the patient has a predisposition to erosion, during the transition from  $t_1$  to  $t_2$  an erosion object is created.
- The patient author predefines the maximum number and size of erosions for the given patient.
- The erosion object changes, as does the patient's symptom profile, over the period of  $t_2$  via interpolation.

# Treatments

- For pre-tumor GERD, treatments are modeled as essentially reversing the disease path at a faster rate than the original progression of the disease.
- For other diseases, like achalasia, treatments can fundamentally change the physiology of a patient – even automatically giving rise to a new disease as a side-effect.

## Side-Effect Diseases: An example

- Achalasia is a disease that makes the lower esophageal sphincter too tight, not allowing food to pass during swallowing.
- Heller Myotomy is a surgical procedure that cuts the lower esophageal sphincter, as a result of which it cannot contract any more.
- By definition, a successful Heller Myotomy turns an achalasia patient – or a healthy person (!) – into a GERD patient.



# Scripts

- We elicit and record knowledge about basic physiological processes as event scripts.
- These scripts are recorded in the **has-event-as-part** property of events in the OntoSem ontology.
- For example, the **swallow** script, which is central to the modeling of diseases of the esophagus, has two subevents, the **oropharyngeal-phase-of-swallow** and the **esophageal-phase-of-swallow**, each of which has its own subevents, which have their own subevents, and so on.
- The subevents in event scripts are instantiated if their preconditions are met.



## An illustration:

An ontological complex event (“script”) and its associated knowledge and rules

swallow	preconditions
	effects
	roles
	composition
	properties

swallow	preconditions	bolus location: mouth
	effects	
	roles	
	composition	
	properties	

swallow	preconditions	
	effects	bolus location: stomach
	roles	
	composition	
	properties	

swallow	preconditions		
	effects		
	roles	agent	human
		theme	bolus
	composition		
	properties		

swallow	preconditions	
	effects	
	roles	
	composition	oropharyngeal phase of swallowing esophageal phase of swallowing
	properties	

swallow	preconditions	
	effects	
	roles	
	composition	
	properties	duration 10.0

oropharyngeal phase of swallowing	preconditions
	effects
	roles
	composition
	properties

oropharyngeal phase of swallowing	preconditions	bolus location: mouth
	effects	
	roles	
	composition	
	properties	



oropharyngeal phase of swallowing	preconditions	
	effects	<b>bolus</b> location: <b>pharynx</b>
	roles	
	composition	
	properties	

oropharyngeal phase of swallowing	preconditions		
	effects		
	roles	agent	human
	composition	theme	bolus
	properties		

oropharyngeal phase of swallowing	preconditions	
	effects	
	roles	
	composition	
	properties	duration 1.0

oropharyngeal phase of swallowing	preconditions	
	effects	
	roles	
	composition	motion event 1 contract muscle 1 motion event 2 relax muscle 1 relax muscle 2
	properties	tongue moves bolus from mouth to pharynx contract pharynx, epiglottis closes bolus moves to larynx, epiglottis opens cricopharyngeus relaxes LES relaxes

(oropharyngeal phase of swallowing)  <b>motion event 1</b>	preconditions	
	effects	<b>bolus</b> location: <b>pharynx</b>
	properties	duration 0.8
	composition	
	roles	agent human theme bolus instrument tongue source mouth destination pharynx

esophageal phase of swallowing	preconditions	
	effects	bolus location: stomach
	roles	agent *none* theme bolus
	properties	duration 9.0
	composition	for $i = L_1$ to $L_{n-1}$ for $j = L_2$ to $L_n$ do peristalsis source $i$ destination $j$

$L = \{ \text{larynx}$   
 $\quad C6$   
 $\quad C7$   
 $\quad T1$   
 $\quad T2$   
 $\quad T3$   
 $\quad T4$   
 $\quad T5$   
 $\quad T6$   
 $\quad T7$   
 $\quad T8$   
 $\quad T9$   
 $\quad T10$   
 $\quad \text{stomach} \}$

esophageal phase of swallowing	preconditions	
	effects	bolus location: stomach
	roles	agent *none* theme bolus
	properties	duration 9.0
	composition	for i = $L_1$ to $L_{n-1}$ for j = $L_2$ to $L_n$ do peristalsis source i destination j

$L = \{ \text{larynx}$   
 $\text{C6}$   
 $\text{C7}$   
 $\text{T1}$   
 $\text{T2}$   
 $\text{T3}$   
 $\text{T4}$   
 $\text{T5}$   
 $\text{T6}$   
 $\text{T7}$   
 $\text{T8}$   
 $\text{T9}$   
 $\text{T10}$   
 $\text{stomach} \}$

peristalsis	preconditions	bolus location $L_i$	
	effects	= effects of motion event of peristalsis	
	roles	agent theme	*none* source $L_i$ bolus destination $L_{i+1}$
	properties	duration	1.0
	composition	stretch fire-nerve stimulate 1 stimulate 2 contract muscle relax muscle 1 motion event relax muscle 2	

lumen of source stretches  
 stretch receptor in source  
 receptor in source stimulates nerve in source  
 receptor in source stimulates nerve in destination  
 source muscles contract  
 destination muscles relax  
 bolus moves to destination  
 source muscles relax



peristalsis	preconditions	bolus location $L_i$	
	effects	= effects of motion event of peristalsis	
	roles	agent *none* source $L_i$ theme bolus destination $L_{i+1}$	
	properties	duration 1.0	
	composition	stretch fire-nerve stimulate 1 stimulate 2 contract muscle relax muscle 1 <b>motion event</b> relax muscle 2	lumen of source stretches stretch receptor in source receptor in source stimulates nerve in source receptor in source stimulates nerve in destination source muscles contract destination muscles relax bolus moves to destination source muscles relax

(peristalsis)  <b>motion event</b>	preconditions	
	effects	<b>bolus</b> location $L_{i+1}$
	properties	<b>bolus solid?</b>
	composition	lumen of $L_i$ $1.4 \Leftrightarrow 1.8?$ add <b>dysphagia</b> intensity $0.1 \Leftrightarrow 0.3$
	roles	lumen of $L_i$ $1.0 \Leftrightarrow 1.39?$ add <b>dysphagia</b> intensity $0.31 \Leftrightarrow 0.69$ lumen of $L_i < 0.99?$ epistemic 0 (bolus stays in $L_i$ )
		<b>bolus liquid?</b>
		lumen of $L_i$ $0.31 \Leftrightarrow 0.99?$ add <b>dysphagia</b> intensity $0.31 \Leftrightarrow 0.69$
		lumen of $L_i < 0.3?$ epistemic 0 (bolus stays in $L_i$ )

# Object-Oriented Scripts

The **swallow** script just described is event-oriented.

We also use object-oriented scripts, which are triggered by changes in the property values of objects.

At each clock cycle, the simulator checks which preconditions for object-oriented scripts have been met and launches those scripts.

For example, the **heartbeat** script is object-oriented: the precondition for the heart contracting and uncontracting is the passing of a certain number of clock cycles.

## Plan of Talk

1. What is the Double Agent System?
2. How it works: a demo
3. Knowledge sources and knowledge elicitation
4. Discussion

## **Some Important Properties of Double Agent**

- Hybrid knowledge
- Hybrid control
- Autonomous eventualities
- Time
- Changes in paradigm of learning
- Double dipping
- Feasibility



# Hybrid knowledge

- a mixture of volitional and non-volitional events
- a mixture of physiological and clinical knowledge (medicine is full of gaps in knowledge; less than 10% of physiological properties are understood; need bridges). says,

“Some computer scientists feel betrayed when they realize how fuzzy the rest of biology [outside of DNA sequencing] is... Biology's theoretical basis is still in its infancy, so few 'first principle' approaches have any chance of working yet”.  
Altman's (2001:14)

# Hybrid Control

- Double Agent models a mixture of volitional and non-volitional events
- and employs a mixture of control structures (object-oriented and event-oriented)

# Autonomous Eventualities

- allows for unexpected input (not canned);
- the user can ask for any test or apply any treatment to any patient at any time, even if medically unsound (and therefore not part of a “best treatment” script)



# Time

- Double Agent supports variable time granularity, from milliseconds (e.g., muscle contraction) to years (e.g., tumor growth)
- discrete modeling at this time...

# The Paradigm of Learning

- Theme-and-variations paradigm permits construction of hundreds of patient instances automatically.
- Permits learning by repetition, extensive practice.
- Permits learning by experimentation (students can even learn about a disease from scratch through trial and error).

# Double Dipping

- Same ontology and processors for NLP and simulation
- Supports adding NLP to the simulation interface
- Eases knowledge bottleneck

# Feasibility Concerns

- Incorporates available high-quality resources, like Foundational Model of Anatomy
- No attempt to recreate a human in the box, but creating a model that behaves like a human in all relevant ways
- Use bridges for as yet unknown physiological properties.

## A few other tasks and approaches in medical simulation and training

1. technical task trainers, which concentrate on a technical task and include only the minimal amount of cognitive simulation necessary for the user to understand a specific technical step, like inserting a needle.
2. non-biomechanistic mannikin trainers (e.g., “SimBaby”, Laerdal, Inc., and “The Human Patient Simulator”, Medical Education Technologies, Inc.), which focus on a narrow scope of acute physiological processes.



3. “canned” scenarios based on clinical decision-making algorithms at the case-level (e.g., MedCases, Inc.); user options are restricted and responses are highly pre-scripted to provide predetermined patient outcomes.
4. Sim-Patient developed by RTI, where acute traumatic patient scenarios are available to a user; the data structures and content of this system are proprietary.
5. Virtual Soldier, which seeks to produce a simulation of the human thorax that is functional for penetrating trauma; focus on anatomical trauma and intelligently guiding emergency interventions.

# CIRCSIM

The goal: to teach about the baroreceptor reflex, the body's rapid response system for dealing with changes in blood pressure.

The original system, MacMan: a mathematical model that could be explored; "naked simulation," not tutoring.

The next system, Heartsim: some feedback but they realized the mathematical model wasn't being exploited.

"The most effective teaching was being generated from the stored correct predictions for each procedure, not from the quantitative outputs generated by the model."

# Appendix 3-C

## Cognitive Simulation in Virtual Patients

**Sergei Nirenburg, Marjorie McShane,  
Stephen Beale, Thomas O'Hara**

University of Maryland, Baltimore County  
{sergei,marge,sbeale,tomohara}@umbc.edu

**Bruce Jarrell, George Fantry, John Raczek**

University of Maryland School of Medicine  
BJarrell@som.umaryland.edu,gfantry@medicine.umaryland.edu  
jraczek@clc.umaryland.edu

### Abstract

We present an overview of the Virtual Patient project at the University of Maryland, which is developing a cognitive model of humans experiencing various states of health and disease to be used in interactive simulations for physician training.

### Overview

This Virtual Patient<sup>1</sup> project is devoted to creating a cognitive, knowledge-based model of a virtual patient (VP) that undergoes both normal and pathological physiological processes. VPs are ontological objects, specifically, subclasses of VIRTUAL-HUMAN that have various diseases and disorders. Like all VIRTUAL-HUMANS, their large inventory of property-value pairs changes in response to ontological events, including internal and external stimuli. All VPs inherit the lion's share of physiology from VIRTUAL-HUMAN, meaning that GERD-PATIENT and HEART-DISEASE-PATIENT (as ontological concepts, not instances) differ only with respect to the disease-specific changes that affect certain of their property values over time.

A cornerstone of a realistic learning environment is creating a wide variety of instances of VPs with a given disease. The basic model of a disease typically involves many tracks (i.e., paths of disease progression), and property values of VPs may fluctuate within specified ranges, adding to the variety of possible VP instances. Authoring an instance of a VP (typically done by a physician-teacher or disease specialist) involves establishing specific values of the VP's basic physiological properties, relevant lifestyle factors, the rate and direction of progression of the disease, the specific symptom profile at given times, and so on. The VP authoring process is, thus, similar to a multiple-choice questionnaire that takes little time to complete. In fact, large inventories of instances of VPs with a particular disease can be generated automatically on the basis of a relatively small inventory of basic ontological "models."

<sup>1</sup> Patent pending.

Importantly, once a VP instance is under the care of a medical student, whatever treatment the student administers causes a change in the state of the VP instance. VP instances are, thus, not canned scenarios, they are flexible software agents; and if a student causes a deterioration in a VP instance's state in some unexpected way, he or she needs to recover from that error by treating the patient in this new condition. What makes modeling VPs feasible – since comprehensively modeling human physiology in the abstract would be a boundless and amorphous endeavor – is our goal-oriented approach: we are not trying to recreate the human organism in all its details, we are trying to recreate it only to the extent necessary to support its autonomous functioning in useful training situations.

### The OntoSem Environment

The main knowledge substrate for the VP project is the OntoSem (Ontological Semantics) ontology (Nirenburg and Raskin 2004). This ontology was initially developed to support knowledge-based language processing, but the rich inventory of features (properties) it uses – including descriptions of complex events (scripts) – has facilitated a smooth transition to broader modeling and simulation applications. In augmenting our general-purpose ontology with knowledge from the medical domain, we are including knowledge from existing resources, notably, the Foundational Model of Anatomy (FMA), whose structure and terminology is becoming the standard in the field.

Using the same environment for medical modeling as for natural language processing has two significant advantages. First, many of the architectural and expressive means apply equally well to both domains: the language-independent, property-rich ontology; the use of scripts; the division of labor between ontology, fact repository (a repository of concept instances) and lexicon; and the management of instances within the fact repository. Second, to be really useful, interactive systems must include natural language communication, and high-quality level natural language processing (NLP) is what OntoSem developers have been pursuing for over a decade. Using the same



**representation language and “world view” to create a simulation and to interact with it promises accuracy and efficiency throughout the system.**

## **Comparisons with Other Systems**

A common type of medical simulation is realized in technical task trainers, which concentrate on a technical task and include only the minimal amount of cognitive simulation necessary for the user to understand a specific technical step, like how to insert a needle. A second type is non-biomechanistic manikin trainers (e.g., “SimBaby”, Laerdal, Inc., and “The Human Patient Simulator”, Medical Education Technologies, Inc.), which focus on a narrow scope of acute physiological processes. A third type covers scenarios based on clinical decision-making algorithms at the case-level (e.g., MedCases, Inc.); in these, user options are restricted and responses are highly pre-scripted to provide predetermined patient outcomes. A more sophisticated type of medical simulation is the Sim-Patient developed by RTI, where acute traumatic patient scenarios are available to a user; however, few details about this system are available as the data structures and content are proprietary.

A well-known project is the Virtual Soldier (<http://www.virtualsoldier.net/>), which seeks to produce a simulation of the human thorax that is functional for penetrating trauma. The Virtual Soldier project differs from the VP project in its focus on anatomical trauma and emergency interventions, as opposed to the diagnosis and long-term care of patients experiencing disease.

Another notable simulation environment is CIRCSIM (Illinois Institute of Technology), which teaches about the baroreceptor reflex, the body’s rapid response system for dealing with changes in blood pressure. The history of this project shows a movement away from research on the mathematical model toward research on pedagogical aspects of online tutoring: “The most effective teaching was being generated from the stored correct predictions for each procedure, not from the quantitative outputs generated by the model” (Michael and Rovick 1996). We plan to incorporate some pedagogically-oriented results from CIRCSIM into the further development of the VP environment, but will take a different approach to the language processing aspect (which that team has deemed integral to teaching systems; Evens et al. 2001), since we already have a rich language processing system in place.

In the past 4 years, the International Meeting on Medical Simulation has produced over 200 papers, none of which describe the type of cognitive simulation being pursued in the VP project. A similar absence of cognitive simulation efforts is reflected in the past four years of the journal *Artificial Intelligence in Medicine*.

In the AI tradition, arguably the most well-known script remains Schank and Abelson’s (1977) restaurant script. Most previous efforts to implement scripts were in a narrow domain, and typically suffered

when some unforeseen script move was encountered.<sup>1</sup> In fact, the difficulty in implementing scripts has led to their relatively marginal status in AI. We believe we can largely circumvent the problem of unexpected input by using an ontological substrate that includes default effects of all events that can affect a virtual patient – be they medical interventions or events of daily life, like smoking. The default effects will be used in all cases when specific, non-default effects of events have not been encoded by the author of the given VP instance. Our team of physicians and knowledge engineers is working to arm the system with sufficient domain knowledge to permit any VP to respond in a reasonable way to any available intervention at any time.

## **The Present and Future of the VP Project**

At present, the ontological substrate of the VP concentrates on the esophagus. It covers the anatomy of the esophagus, the physiology of swallowing (including esophageal peristalsis), and a selection of pathologies (such diseases of the esophagus as achalasia, GERD and esophageal tumors). Whenever medical knowledge allows, diseases and diagnostic and treatment procedures are described through their mechanisms and in terms of accessing or modifying the ontological properties of various anatomical elements. In other cases, bridging is used, which can be described as using temporal chains as the substrate for scripts instead of causal chains, pending the discovery of the latter.

In addition to the ontological work, the project is also developing a simulation environment for the VP, an interaction environment between the VP and the operator, and a library of VP instances with specific diseases to support simulation and training. Results of that development and experimentation will be reported separately. We plan to extend the coverage of the VP to the entire gastrointestinal tract and then beyond it to other systems in the organism. We are also working on a mentoring component for the system and on natural language interaction capabilities for the VP.

## **References**

- Clark, P.; and Porter, B. KM: The Knowledge Machine 2.0: Users Manual. <http://www.cs.utexas.edu/users/mfkb/km.html>.
- Cullingford, R. 1981. “SAM”. *Inside Computer Understanding*, eds. R. C. Schank and C. K. Riesbeck. Hillsdale, N.J.: Erlbaum.
- Evens, M.; Brandle, S.; Chang, R.; Freedman, R.; Glass, M.; Lee, Y.; Shim, L.; Woo, C.; Zhang, Y.; Zhou, Y.; Michael, J.; and Rovick, A. 2001. CIRCSIM-Tutor: An Intelligent Tutoring System Using Natural Language Dialogue. In *Proceedings of the 12th Midwest AI and Cognitive Science Conference*, Oxford OH, 2001, pp. 16–23.

---

<sup>1</sup> Cullingford et al.’s SAM system (Cullingford 1981) is among the most well-known. It read stories that corresponded to scripts and output multi-lingual summaries and answered questions about the texts.

Michael, J.; and Rovick, A. 1996. The Once And For All History of HEARTSIM/CIRCSIM.  
<http://www.cs.iit.edu/~cirsim/prehistory.html>.

Nirenburg, S.; and Raskin, V. 2004. *Ontological Semantics*. Cambridge, Mass.: MIT Press.

Schank, R.; and Abelson, R. 1977. *Scripts, Plans, Goals and Understanding*. Hillsdale, NJ: Earlbaum Assoc.

## **Exhibit 3-D**

# **DRAFT: Training Clinical Decision Making Using Cognitively Modeled Virtual Patients**

Marjorie McShane\*, Bruce Jarrell<sup>+</sup>, Sergei Nirenburg\*, George Fantry<sup>+</sup>, Stephen Beale\*

\*University of Maryland Baltimore County

<sup>+</sup>Univeristy of Maryland Medical School

## **1. Introduction**

Clinical decision making skills are developed through practice. One drawback of current physician training methods is a lack of sufficient training scenarios, in live or simulated patients, to support a student's mastering of the complex web of biomechanical and clinical knowledge used daily by practicing physicians. Ideally, when studying a disease, students should have access to a population of patients suffering from that disease, with each patient displaying clinically relevant variations on the theme. Such variations might involve the path or speed of disease progression, the profile and severity of symptoms, responses to treatments, and secondary diseases or disorders that affect treatment choices. If each student could learn by independently managing the care of many such patients – especially in a context in which trial and error learning carried no risk – we would expect their decision making skills to develop faster than with traditional training methods alone. We set out to create such a learning environment. We specified that the environment would need to allow student management of a variety of disease categories including chronic and acute disorders as well as simple and complex diseases. The environment would also need to accommodate both poorly and well-defined knowledge about disease processes, and clinically relevant variations, such as complications of the disease, complications from associated treatment modalities and time course variations. Accomplishing this simulation creates one component of the learning environment – the

patient – and the opportunity for trial and error management by a student. It does not address the second component of the environment, the task of mentoring the student. A mentoring component adds yet another layer of complexity. As one can see, the overall task is quite challenging, but the benefit to the learning environment would be huge if successfully accomplished.

The Maryland Virtual Patient (MVP) project<sup>1</sup> seeks to address these tasks by developing autonomous, flexible computer simulations of patients coupled with a virtual tutor. The simulated patient and the tutor are capable of interacting with students using natural language. What we will describe has been created in prototype and has resulted in a “living” simulated patient. In this overview of the MVP project, we describe our knowledge-based, ontologically grounded approach to modeling and simulation and how this approach has permitted us to handle complexity by generating automatic functions and responses in virtual patients of the Maryland pedigree (MVPs). As designed, MVPs represent a conceptual leap in the computer modeling of humans in the continuum between health and disease. They embody and make manifest biophysical functions that have clinical relevance in the maintenance of health, the production of disease, and the bidirectional transitions between these two states. They are autonomous, their state evolving over time and in response to both internal and external stimuli.

The knowledge encoded in the MVP mirrors that used daily by clinicians managing patients and mentoring students. It includes structure and function information derived from well-understood anatomy and physiology literature and ranges from the molecular level up to the level of organism. Where gaps exist in the knowledge of biomechanisms, the knowledge is bridged with practical clinical knowledge and observations derived from the literature. This hybrid knowledge base reflects precisely what a clinician employs when working with a patient. The level of granularity for modeling both types of mechanisms

---

<sup>1</sup> Patent pending.

is set by the requirements of the task-oriented simulation. By limiting descriptive granularity to that needed for simulation, we are not required to include every mechanism known to biology and clinical medicine, only a practical subset. It is this circumscription of the task that renders MVP modeling feasible.

A key design feature of the MVP is automatic function in response to internal stimuli and external interventions. As an example of automatic function in response to external interventions, the MVP is able to respond realistically to students' questions or to any treatments (from among those available in the system at a given time) prescribed by the student. Response to a question is a cognitive operation while response to treatment is a physiological operation; both kinds of simulation are supported in the MVP. If the student launches an inappropriate treatment, the MVP's state will not improve and may even deteriorate, in which case the student must recover from his mistake. The system need not exhaustively list all permutations of paths a student could take and all consequential responses of the MVP; instead, it relies on ontologically-grounded descriptions of basic physiology, disease processes, effects of treatments, and so on, so that the state of a given MVP at a given time will, quite literally, be composed by the underlying model. Automaticity has important consequences for creating individual MVPs with a disease because it limits the number of variables an author has to manipulate to create variations on the basic model.

As just framed, this task might sound daunting. In fact, the developers of one advanced medical tutoring system, CIRCSIM-Tutor, seem pessimistic about the prospects of automatic tutoring in a less than highly constrained realm:

“When we started the CIRCSIM-Tutor project 15 years ago, some experts in the field argued that student modeling was too difficult to be worth the trouble; some even classified the problem as totally intractable... Anyone who observes human tutors in action, on the other hand, must

recognize that they base decisions at all levels, from the choice of the next problem to present to the student to what kind of hint to provide, on their model of the student... Joel Michael and Allen Rovick were so convinced of the crucial importance of modeling that they picked the CIRCSIM domain [the baroreceptor reflex] for our tutor largely because they felt that it would be easy to construct a good student model in this subject area ... They are... convinced that it is important to build a comprehensive model before starting to tutor, to ensure that the tutor can begin by attacking the most important of the student's conceptual difficulties” (p. Evens and Michel 2006: 252-3)<sup>1</sup>

Clearly, selecting a narrow domain facilitates domain modeling, student modeling and the automation of tutoring support; and, all other things being equal, one would expect better near-term results from a more highly constrained system. However, all other things never are really equal: real-world needs can also set the agenda for research and development. We have a predefined task for which a sufficient, real-time solution must be found; and while task-driven projects necessarily involve unknowns, they also promise exciting new horizons both within the targeted application and beyond it.

## 2. Modeling and the MVP

An MVP is a computer-based model of human physiology grounded in a formal ontology, or world model. MVPs are specialized classes of Virtual Humans, meaning that they are Virtual Humans suffering from a particular disease or disorder. As such, MVPs with achalasia and MVPs with GERD

---

<sup>1</sup> This book provides a useful review (citing on the order of 600 references) of tutoring systems, the tutoring literature, natural language processing as applied to tutoring, and related topics.

have the lion’s share of physiology in common, apart from disease-specific anomalies.

The ontology that underpins the MVP system – called the OntoSem ontology – is quite different from most other ontologies by virtue of its rich property-based descriptions of concepts and its coverage of three types of entities: objects, events and properties.<sup>1</sup> By contrast, most ontologies (e.g., UMLS, Bodenreider 2004) are hierarchical word nets that contain few or no properties; many exclude events entirely. For a taste of the structure and content of the OntoSem ontology, see Figure 1, which is a screen shot of the concept ESOPHAGUS. The right-hand panel shows part of the frame for ESOPHAGUS. As one can see from the slider for the right frame, only a small portion of the inherited properties can be shown on the screen at one time. Property values are locally specified for a concepts in need-based fashion, since locally specifying the maximum number of property values for each concept would require extensive resources and is often not required for our application-oriented work. The left-hand panel shows where ESOPHAGUS is located in the IS-A hierarchy of the ontology.

<sup>1</sup> The OntoSem ontology derives from the theory of Ontological Semantics, a theory originally developed for knowledge-rich text processing (Nirenburg and Rakin 2004). It should be mentioned that, while most ontologies have not proven useful for our work, one has: the Foundational Model of Anatomy (FMA) (Rosse and Mejino 2004; <http://fma.biostr.washington.edu>). FMA provides both inheritance (is-a) and meronymic (part-of) trees for elements of human anatomy. Concepts are linked using a mid-sized inventory of properties. As we are supplementing the OntoSem ontology for use in the medical domain, we are following the FMA model in certain ways (e.g., with regard to naming conventions) in order to keep our knowledge resources compatible with what we believe will become the accepted standard. However, it would be incorrect to assume that FMA answers all our needs in the medical domain: it treats only anatomical objects, whereas we need a full treatment of relevant events and their relationship to objects, both anatomical and extra-anatomical.

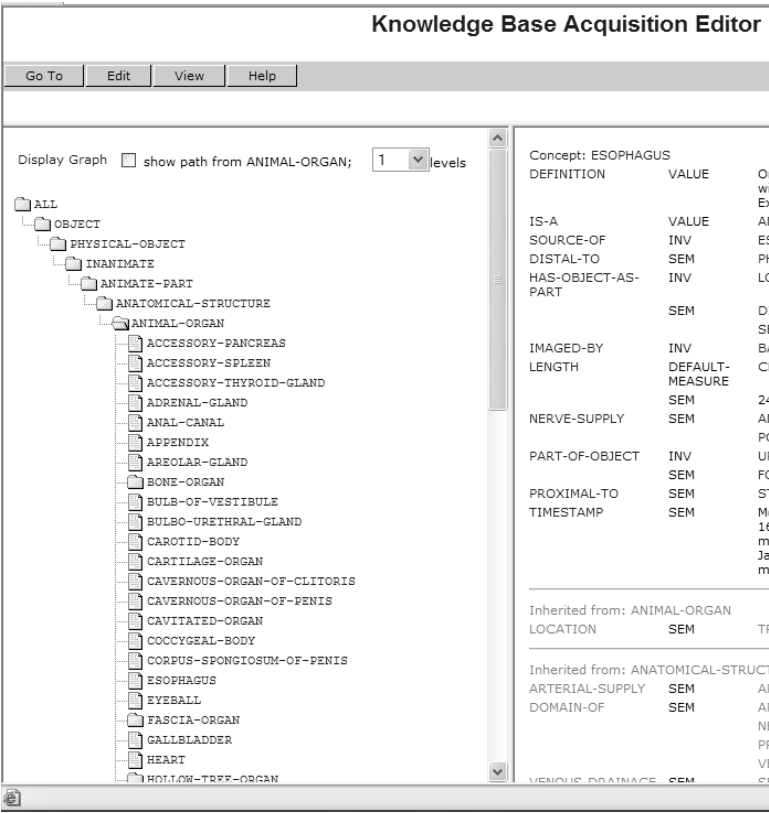


Figure 1. The concept ESOPHAGUS in the OntoSem ontology.

Crucially for the MVP project, the metalanguage of ontological description in OntoSem supports the encoding of complex events, also known as scripts. Scripts represent typical sequences of events and their causal and temporal relationships. In other words, they encode how individual events hold well-defined places in routine, typical sequences of events that happen in the world, with a well-specified set of objects filling different roles throughout that sequence. For example, if the event is swallowing, there is only one animate participant (the swallower), but many other objects play necessary roles: various nerves and muscles act as instruments of peristalsis; the swallowed bolus is the theme of peristalsis-driven motion events; the stomach is the final destination of the bolus, and so on. Scripts normally contain subscripts and can be more or less fine-grained depending on the goals of the given simulation.

Within the MVP project we are developing both *domain* scripts and *workflow* scripts. Domain scripts describe basic physiology, disease progression and responses to treatments, whereas workflow scripts model the way an expert physician would handle a case, thus forming the knowledge substrate for automatic tutoring. It must be emphasized that although the scripts described below do not correspond to the frame-based format used in the interface for the core ontology (Figure 1), they are actually a full-fledged part of that ontology (i.e., unified world model).

Each *instance* of an MVP – for example, Barry Hume who is suffering from fast-progressing achalasia – is an inventory of property values stored in a dynamic database. The properties of MVPs change over time in response to both internal stimuli (e.g., the necessity of the heart to beat rhythmically; the progression of a disease) and external stimuli (e.g., eating chocolate; taking a drug; having surgery). The database that stores information about MVP instances is called the fact repository. It is linked to the ontology and employs the same metalanguage of description. So, just as the relationships between *types* of objects and events are stored in the ontology (the descriptive component of the knowledge base), the relationships between *instances* of objects and events are stored in the fact repository (the assertion component of the knowledge base).

A cornerstone of creating a realistic MVP environment is providing for wide variation among instances of MVPs with a given disease. That is, the basic model of a disease includes all relevant tracks (i.e., paths of progression), and each track provides many choice points that differentiate cases. The author of a given instance of a MVP (typically a physician-teacher or disease specialist) then determines that MVP's basic physiological properties, relevant lifestyle factors, the rate of progression of the disease, which path the disease takes at all possible furcations, the specific symptom profile and given times, and so on. The MVP authoring process is essentially a multiple-choice questionnaire that takes little time to complete. The reason why VPs can so readily be authored derives

from the care taken to create the basic model of the disease, including delineating exactly which property values are available for individual parameterization and which ones are fixed for all patients experiencing the given stage of a disease.

A conceptual paradigm that has proved useful for modeling physiology and disease progression was laid down by Altman, who recognizes "a need for cross-scale data integration methods" for intelligent systems in biology (Altman et al. 2001: 14). Altman focuses on integrative modeling of biomechanisms spanning gene and protein levels, which is necessary to automatically produce complex protein structure and function. In his model, a group of genes and initial conditions are the starting point for building a protein. This starting point can then be chained through biomechanistic computational pathways to automatically produce a highly specified functional protein. If the starting point is a mutated gene or abnormal initial condition, then the model will produce an abnormal protein. In short, Altman starts from first principles to derive a complex event. One insufficiency of the Altman model, however, is that it does not incorporate environmental conditions that are also believed to affect the genesis of a protein, since those influences are poorly understood and cannot be incorporated into a strictly first-principles approach.

The model we are building takes a mixed approach: *first* principles plus *clinical* principles. Modeling using first principles is undertaken when the relevant physiological pathways are known and when the modeling of those pathways is believed to have pedagogical utility; otherwise, clinical knowledge is used as a "bridge" in the simulation. Looking past the immediate educational application of the VP, we envision a complementary line of work that pursues the modeling of biomechanistic pathways to make the functioning of the VP ever more dependent upon causal chains and ever more true to the actual functioning of a human being.

It is fair to say that the VP project redefines descriptive adequacy for (a) medical mechanisms and their interactions from the level of gene to the level of population, and (b) clinical knowledge and interventions for the diagnosis and treatment of

disease. It also promises progress in knowledge representation strategies, as it requires the modeling of objects and events that change over time both in response to internal stimuli and external intervention.

The MVP project places significant demands on physician-informants to render complex, multi-scale functions in a form that can be implemented computationally – naturally, with a knowledge engineer mediating between physician and programmer.<sup>1</sup> Physicians must distill their extensive and tightly coupled physiological and clinical knowledge into the most relevant subset, and express it in the most concrete of terms. Not infrequently, they are also called upon to hypothesize about the unknowable, like the state of a patient experiencing a pre-clinical stage of disease, or the state of a patient after an effective treatment that is never, in real life, followed up by objective tests. Such hypotheses reflect the mental models of given experts, which might differ in subtle ways from those of other experts. However, such differences, we would suggest, have little bearing on the ultimate goal of this enterprise: to create MVPs whose behavior is sufficiently life-like to further specific teaching goals.

### 3. Modeling Diseases, Symptoms and Treatments

In the MVP system, diseases are modeled as changes in key property values of an MVP over time. For each disease, a set number of stages is established, and typical values (or ranges of values) for each property are associated with each stage. Values at the start of each stage are recorded explicitly, with values between stages being interpolated (the interpolation currently uses a linear function, though other functions could as easily be employed). The disease model includes a combination of fixed and variable features. For example, although the number of stages for a given disease is fixed, the duration of

<sup>1</sup> We plan to support this process using a smart computer interface, but do not realistically expect the mediation of knowledge engineers to be expendable in the near term.

each stage is variable; similarly, although the values for *some* physiological properties undergo fixed changes across patients, the values for other physiological properties are variable within a specified range. The combination of fixed and variable features represents, we believe, the golden mean for disease modeling. On the one hand, each disease model is sufficiently constrained so that MVPs suffering from the disease must show appropriate physiological manifestations of it. On the other hand, each disease model is sufficiently flexible to permit instances of MVPs to differ in clinically relevant ways, as selected by the author of each MVP.

To concretize the discussion of disease modeling, let us consider the example of achalasia, a disease that progressively renders a patient unable to swallow, which is thought to be due to a loss of relaxing neurons in the lower esophageal sphincter (LES). Table 1 encapsulates the generic physiological profile of VPs with achalasia. Expressive means used in the table are described in the explanatory text that follows.

**Table 1: Physiological properties that change due to achalasia**

	<i>disease stages (default duration in months)</i>	<b>t0 (6)</b>	<b>t1 (6)</b>	<b>t2 (6)</b>
1.1	Ratio of contracting to relaxing neurons in the LES	100/100	75/100	50/100
1.2	Basal LES pressure (mmHg)	<b>0 - 40</b> (25)	$P_{t0} + 10$	$P_{t0} + 20$
1.3	Residual LES pressure (mmHg)	0	10	20
1.4	Relaxed LES diameter (cm.)	2	1.5	1.0
1.5	amplitude of contraction during peristalsis	80	60	35
1.6	efficacy of peristalsis (reflected in changes during swallowing)	<i>normal</i>	<i>normal</i>	<i>intermittent peristalsis</i>
1.7	diameter of distal	2	3	4



	esophagus (cm.)			length of	time it takes to swallow a bolus increases
1.8	bolus quantity in distal esophagus (on abstract scale 0-1)	0	0 - .2 (.1)	.2 - t4, essentially, “never”, which is arbitrarily encoded as 35,000 minutes. Although the quantity of accumulated boluses (row 1.8) and the duration of swallowing (row 1.9) are “basic” physiological properties of an MVP, they are objectively	
1.9	duration of swallowing (min.)	0	1	5	

Achalasia is modeled in 5 stages, t0 – t4 (“time 0 to time 4”), whose default values, in months, are listed in parentheses in the top row of Table 1. The independent variable (row 1.1) is the ratio of contracting to relaxing neurons in the LES, which decreases steadily over the five stages of the disease. All other physiological properties depend upon this variable. The heavy line between the independent variable and the dependent variables underscores this distinction in kind.

The basal LES pressure (row 1.2) at t0 (the pre-clinical stage of the disease) is a variable, ranging from 0 to 40 mmHg, with a default of 25 mmHg (written in parentheses). The author of each MVP instance must select a basal LES pressure for that patient or accept the default. The basal LES pressure increases by 10 mmHg with each stage of the disease ( $P_{t0}$  is the pressure at time 0, which is the basis for subsequent calculations). As such, a VP with a basal LES pressure of 20 at t0 will have a basal LES pressure of 30 at the start of stage t1, 35 in the middle of t1, 40 at the start of t2, etc. The residual LES pressure (row 1.3) increases at a similar rate. The diameter of the LES during relaxation (row 1.4) decreases over time, so that by the time the residual LES pressure is 40, the diameter of the LES during relaxation is 0 cm.; at this point, the LES does not open at all to let food pass.

The efficacy of peristalsis (row 1.6) reduces over the course of the disease due to changes in the ratio of contracting to relaxing neurons in the body of the esophagus. Impaired function is detectable in the t2 stage, escalating to aperistalsis in t3 and beyond.

The quantity of unpassed boluses that accumulate in the distal esophagus (row 1.8) increases over time, as does the diameter of the distal esophagus (row 1.7). (There is an inverse correlation between esophageal diameter and efficacy of peristalsis.) The

The disease table for achalasia contains relatively few variables, highlighted with boldface: the duration of each stage of the disease, the MVP’s basal LES pressure at t0, and the quantity of bolus matter in the distal esophagus at all stages. The reason for having so few parameterizable features is the conviction, on the part of the physicians building the model, that having more parameterizable features would not serve any pedagogical goal. The more useful locus of parameterization among MVPs with achalasia lies in their symptom profiles and their responses to treatment.

The experiencing of symptoms varies widely across patients and, accordingly, cannot be directly linked to given physiological states. However, a fixed inventory of symptoms is associated with each disease, and expected ranges of values for each symptom are asserted for each stage. Defaults are provided as well.

Table 2 shows the symptom profile table for achalasia. The disease is associated with four canonical symptoms, all of which are expected to be experienced to some extent by the advanced stages of the disease. Considering the wide range of values for these symptoms, one can create instances of MVPs suffering from achalasia that present with very different symptom profiles (as in all tables, parameterizable features are highlighted with boldface).

**Table 2: Symptom profile of achalasia**

		t0	t1	t2
2.1	Difficulty swallowing	0	<b>0 - .2</b> (.1)	<b>.1 - .4</b> (.2)
2.2	Weight loss	0	<b>0 - .05</b> (0)	<b>0 - .1</b> (0)
2.3	Chest pain	0	<b>0 - .3</b> (.1)	<b>0 - .5</b> (.3)

2.4	Regurgitation (times/month)	0	0 – 4 (0)	0 – 20 (10)	educational, family oriented, and to have (40) vastly different (76) symptom profiles at different times given the same underlying physiological state.
-----	--------------------------------	---	--------------	----------------	---

In modeling the effects of treatments, close attention is paid to the distinction between changes in property values that need to be asserted and changes that can be inferred by the system. Continuing with the example of achalasia, we know from the disease model that the independent variable is the ratio of relaxing to contracting neurons in the LES; therefore, it would be natural – from a purely computational perspective – to model treatment effects around that variable. However, in our complex cooperation between people and machines, people’s preferences commonly hold sway. A physician’s model of achalasia revolves not around neuron ratios but around basal LES pressure and its implications on the efficacy of swallowing. Therefore, to make the data as transparent as possible to physicians and students, we orient the discussion of treatment effects around the basal LES pressure—from which, of course, the system can infer the ratio of relaxing to contracting neurons and its other dependent property values .

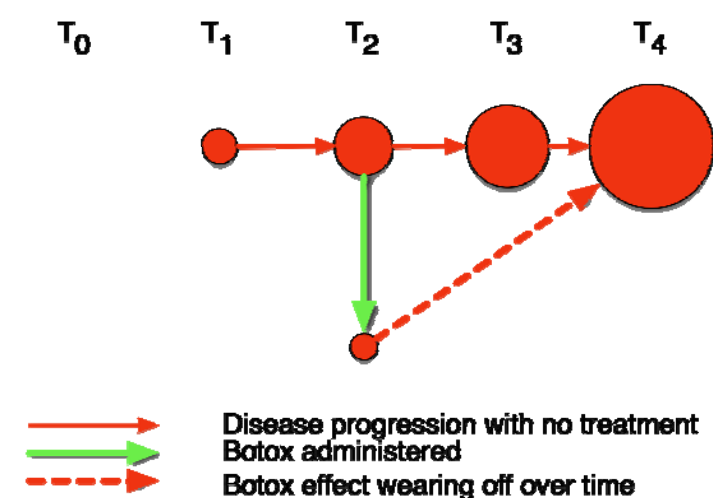
A treatment for achalasia can be effective or ineffective; if effective, its efficacy can be short-term or long-term. Efficacy is judged by the basal LES pressure: if basal LES pressure goes down a significant amount the treatment is effective; if not, it is not.

Most physiological property values and all symptoms remain in a constant relationship with basal LES pressure. As such, if an MVP has a basal LES pressure of 25 mmHg early in his disease, and returns to the same pressure later in his disease after some successful intervention, his overall state should be largely the same at those two points. As such, one does not need to redefine the MVP from scratch in anticipation of each intervention; instead, the original disease and symptom tables can continue to be consulted by the system throughout the MVP’s lifetime. This reuse of the original disease and symptom tables is not only a convenient shorthand, it ensures consistency of the MVPs profile throughout its existence. It would be unrealistic, or at least highly unusual and unnecessary for

We will further pursue the issue of what treatment effects need to be recorded explicitly in describing three typical treatments for achalasia: the administration of BoTox, pneumatic dilation and Heller Myotomy.

**BoTox.** BoTox is a neurotoxin derived from the bacteria, *Clostridia botulinum*. When injected into muscle, the toxin inhibits release of acetylcholine from presynaptic neurons preventing the nerve impulse from reaching the muscle which results in muscle relaxation or paralysis. The net effect in achalasia is to improve the balance between inhibitory and stimulatory innervation of the lower esophageal sphincter resulting in a decrease in basal or resting tone of the lower esophageal sphincter. If BoTox is effective, its effects typically last for 6-12 months. In the current implementation, the durations are exactly 0, 6 and 12 months, but other durations could be included as well if deemed clinically and pedagogically relevant.

BoTox does not cure achalasia, nor does it stop the progression of the disease, it only temporarily alleviates symptoms. Therefore, when its effects wear off, the disease will be at the same stage as it would have been if BoTox had not been administered. Figure 2 is a visual representation of this process.



**Circle size reflects the severity of symptoms**  
**Figure 2. The effect of BoTox on a patient's symptoms of achalasia.**

The circles represent the severity of achalasia symptoms: the larger the circle, the more symptoms experienced. This figure represents the situation in which BoTox was given at the beginning of  $t_2$  and wore off by the beginning of  $t_4$ . The rate of deterioration of the MVP's symptom profile between  $t_2$ -with-BoTox and  $t_4$  is faster than if no treatment had been given (recall that we use a linear interpolation of property values between disease stages, so a larger distance between start and end values over the same amount of time implies a greater rate of change). However, for at least the initial period of treatment, this patient will benefit from having been administered BoTox.

Patients' responses to BoTox vary widely, based in part on the stage of the disease at which the treatment is administered. As such, the author of each MVP instance must select how that patient will respond to BoTox if it should be administered at each stage of the disease. As always, defaults are provided.

**Table 3. The effects of administering BoTox to an MVP at different stages of achalasia**

If a patient is given BoTox during this stage of the disease,	$t_0$	$t_1$	$t_2$	$t_3$	$t_4$
his basal LES pressure will	4 - 24	12 - 30	18 - 36	24 - 42	30 - 48

initially go down to	(15)	(21)	(27)
his residual LES pressure will initially go down to	0 - 5 (0)	3 - 10 (5)	5 - 12 (9)
the effect will wear off over # months	0, 6, 12 (6)	0, 6, 12 (6)	0, 6, 12 (6)
his retained esophageal content will decrease to	0 (0)	0 - .15 (.1)	.1 - .3 (.2)

Of the nine properties in the achalasia disease table, only 3 are asserted in the BoTox treatment table: basal LES pressure, residual LES pressure and esophageal contents. The basal LES pressure must be asserted since that is the one clinically used by physicians as their point of orientation: how an MVP's basal LES pressure changes due to treatment defines how effective the treatment is. The residual LES pressure must also be asserted because the ratio of basal to residual LES pressure is not necessarily the same when a patient suffers from achalasia as when he was healthy. Similarly, either esophageal contents or emptying delay must be asserted, with the second being calculable from the first. The rest of the physiological properties and all of the patient's symptoms can be inferred by the system based on the original correlations in the disease and symptoms tables. The only exception is the circumference of the distal esophagus, which will never return to its original size – a detail that is handled by a special rule in the simulator.

The BoTox table records only the **initial effects** of BoTox therapy on an MVP at different stages of achalasia, since the progression over time is interpolated using the original disease table. As such, the BoTox table should not be read from left-to-right, but only from top-to-bottom, as implied by the wide black borders separating columns.

Let us recap, using the example of BoTox, the earlier description of how property values are calculated during treatments. The simulator can calculate what the disease state would be at any point during a treatment based on two types of information: property values at BoTox-Begin (in the BoTox table) and property values at BoTox-End (the point in the disease where the MVP would have been if no BoTox had been administered). Values

for BoTox-End are drawn from the original disease table. A linear change in values from BoTox-Begin to BoTox-End is assumed. This calculated change in the MVPs state *is* the treatment script – which is different for each MVP instance.

*Pneumatic Dilation.* Pneumatic Dilation (PD) is an endoscopic procedure by which a balloon is inserted into the LES and inflated to rip the muscle layer. PD tends not to reduce basal LES pressure to 0, as a completely successful Heller Myotomy can; instead, a resulting basal pressure of 10-12 mmHg is considered a successful outcome.

There are 3 clinically relevant scenarios representing the efficacy of PD. All are possible regardless of the disease stage at which PD is carried out:

- 1) definitive cure;
- 2) success but regression;
- 3) failure.

Moderate success is not an interesting outcome, pedagogically speaking, and is therefore not included in the model.

We model the results of PD based on clinical knowledge of patient symptoms over time since no tests are done following the procedure to provide objective clinical data on outcomes. The time frames of interest for post-PD patients are preset to the clinically relevant intervals of 1 month, 1 year and 5 years. If the PD fails for a given patient, all of his property values stay as they were, as does the state of his disease.

As with all treatments, the MVP author must decide how his patient will respond to PD if a user should choose to administer it. The first choice is which of the original three scenarios will be selected. If “cure” or “success with regression” is chosen, the MVP author needs to select certain post-PD property values for his patient or else accept the defaults provided in the model.

**Table 4. Pneumatic dilation “cure”**

	Time of PD ( $t_{PD}$ )
Basal LES pressure	<b>10 – 12</b> (10)
Residual LES pressure	<b>0 – 1.5</b> (.5)
Esophageal Contents	<b>0 - .2</b> (.1)
Emptying Delay (mins.)	<b>0 – 5</b> (1)

**Table 5. Pneumatic Dilation “success but regression”**

	Time of PD ( $t_{PD}$ )	$t_{PD} + 1$ mo.	$t_{PD} + 1$ yr.	$t_{PD} + 5$ yrs.
Basal LES	<b>10 – 12</b> (10)	<b>11 – 15</b> (15)	<b>30 – 40</b> (30)	<b>55 – 75</b> (55)
Residual LES	<b>0 – 5</b> (3)	<b>5 – 10</b> (8)	<b>20 – 30</b> (20)	<b>40</b>
Esophageal contents	<b>0 - .2</b> (.1)	<b>0 - .2</b> (.2)	<b>.6</b>	<b>&gt; .7</b>
Emptying delay (mins)	<b>0 – 5</b> (1)	<b>0 – 5</b> (3)	<b>10</b>	<b>35000</b>

*Heller Myotomy.* Heller Myotomy is surgery that cuts rather than tears the LES, potentially resulting in a basal LES pressure of 0 mmHg (a “complete Heller”). The outcome scenarios for Heller Myotomy closely parallel those for pneumatic dilation, apart from different values in some cells of the treatment tables.

**Table 6. Heller Myotomy “cure”**

<i>HM Scenario 1 Table</i>	time of HM
Basal LES pressure	<b>0 – 10</b> (3)
Residual LES pressure	<b>0 – 5</b> (1)
Esophageal Contents	<b>0 - .2</b> (.1)
Emptying Delay (mins.)	<b>0 – 5</b> (1)

**Table 7. Heller Myotomy “success but regression”**

<i>HM Scenario 2 Table</i>	time of HM ( $t_{HM}$ )	$t_{HM} + 1$ mo.	$t_{HM} + 1$ yr.	$t_{HM} + 5$ yrs.
Basal LES	<b>0 – 10</b> (3)	<b>2 – 15</b> (5)	<b>25 – 40</b> (25)	<b>40 – 65</b> (45)
Residual LES	<b>0 – 5</b> (1)	<b>5 – 10</b> (5)	<b>15 – 25</b> (15)	<b>40</b>
Esophageal contents	<b>0 - .2</b> (.1)	<b>0 - .2</b> (.1)	<b>.3 - .6</b> (.3)	<b>&gt; .7</b>
Emptying	<b>0 – 5</b>	<b>0 – 5</b>	<b>6 – 10</b>	<b>35000</b>

delay (mins)	(1)	(2)	(7)	
--------------	-----	-----	-----	--

An interesting aspect of Heller Myotomy is that a full Heller by definition turns an achalasia patient into a GERD patient. Our current model of GERD readily handles this eventuality, since any VP with a basal LES pressure of less than 10 mmHg will experience GERD symptoms. This raises a noteworthy point about treatment modeling in the VP system: all treatments have default outcomes, meaning that if a student performs a Heller Myotomy on a patient with GERD, that patient's GERD will automatically get worse, since his LES pressure will go down – by default – to close to 0. To summarize, MVP authors typically create MVPs with at least some *non*-default physiological properties, symptoms, or responses to treatments. Anything not explicitly selected by MVP authors is handled by defaults recorded in the knowledge base. As such, the system can reason about combinations of eventualities never anticipated by developers. In this way, MVPs behave like animate beings.

The only piece still missing from disease description is how the results of diagnostics are generated. These do not need to be parameterized for individual patients – they derive directly from the property values of the given MVP at the given time as stored in the fact repository.

For each disease covered by the MVP system, all relevant diagnostic procedures are recorded in the ontology. Their ontological description includes the physiological property values that give rise to each positive result. Using this information, the simulator can automatically generate test results for any MVP at any time. Test results can even be correctly generated if an irrelevant test is launched on an MVP: since the patient's property values will typically not match any of those listed as positive results of the test, the result will default to negative – i.e., normal.

For purposes of orientation, we present a tabular overview of tests that are relevant for achalasia and preconditions for returning each possible result. This overview is, essentially, a crib for developers and does not reflect the more elaborate knowledge structures, encoded in the

ontological metalanguage, used by the simulator (we omit these from the table, as they add nothing conceptually to the discussion). The description in the third column is what is returned to a user when the given test is ordered during simulation.

**Table 8. Test Results for Achalasia**

Test	Thumbnail sketches of the physiological preconditions for returning each test result	Test results written using accepted terminology
EGD	LES diameter < .5 cm.	“narrowing of the LES with a pop upon entering the stomach”
barium swallow; EGD	quantity of boluses in distal esophagus > .4 (on an abstract scale of {0 1})	“retained debris”
barium swallow	distal esophagus > 3 cm.	“dilated esophagus”
barium swallow	diameter of LES during swallowing < .5	“bird's beak”
manometry	LES pressure at rest > 45 mmHg	“hypertensive LES”
manometry	LES pressure at rest < 35 45 mmHg	“high-normal LES pressure”
manometry	LES pressure > 8 mmHg during swallowing	“incomplete relaxation of LES”
manometry	peristalsis (epistemic 0)	“aperistalsis”
manometry	peristalsis (epiteuctic .5)	“intermittent peristalsis”
barium swallow	swallow (duration (> 1) (measured-in minute)))	“delayed emptying”

To recap this section, we have described the general approach to modeling virtual patients and diseases within the MVP environment using the example of achalasia to ground the discussion. We have shown how basic disease models incorporate

parameterizable features such that different instances of MVPs can show different, clinically relevant disease paths and outcomes. We have also shown that using an ontological substrate permits the MVP simulator to handle “unexpected” outcomes, which should be a common eventuality in a tutoring environment.

As we hope has become clear by now, one creates instances of MVPs by selecting actual values from among the ranges provided for all parameterizable features in Tables 1-7. Currently, these tables are presented to MVP authors with minimal surrounding text since all authors to date have also been developers (if authors were not developers, then the tables could be embedded in explanatory text similar to that above). To provide a better idea of the actual authoring experience, as well as a glimpse of another disease modeled in the MVP environment, in section 4 we provide the patient-authoring questionnaire for GERD. We trust that readers are now familiar enough with the framework and expressive means of MVP modeling to follow this example with little point-by-point guidance.

## **4 The Patient-Authoring Questionnaire for GERD**

The GERD-patient survey is composed of two parts: first, the author provides basic information about the patient and selects one of seven basic disease tracks for his patient. Then, based on the disease track chosen, he provides follow-up information. We present the survey in a different font to distinguish it from the running text, and use certain shorthands for reasons of space. The descriptive text within the survey is intended to support the work of MVP authors, be they developers or not.

*Start of survey*



## Authoring a GERD Patient, Part I (for all GERD patients)

Patient name \_\_\_\_\_, age \_\_\_\_\_, gender \_\_\_\_\_, race \_\_\_\_\_, basal LES pressure \_\_\_\_\_, basal gastric pressure \_\_\_\_\_.

Does the patient experience transient LES relaxations? (y/n)

Which if any substances, when ingested by the patient, cause GERD symptoms (chocolate, caffeine, mints, alcohol, fatty food, alcohol, a large meal)? \_\_\_\_\_

There are 3 meta-scenarios for GERD that divide into 7 actual subtypes:

1. inflammation – erosion – ulcer – peptic stricture
  - a. inflammation
  - b. inflammation – erosion
  - c. inflammation – erosion – ulcer
  - d. inflammation – erosion – ulcer – peptic stricture
2. inflammation – Barrett's esophagus – tumor
  - a. inflammation – Barrett's esophagus
  - b. inflammation – Barrett's esophagus - tumor
3. proximal GERD

The scenario for a given patient reflects his predispositions. For example, if a patient experiences scenario 1b, it means that, for him, the properties "predisposition-to-inflammation" and "predisposition-to-erosion" have a value of "yes", and the properties related to all other possible GERD predispositions have a value of "no".

Which GERD profile does your patient have (1a, 1b, 1c, 1d, 2a, 2b, or 3)? \_\_\_\_\_

*Let us assume, for this example, that the VP author selects scenario 1c.*

### Scenario 1c ("GERD to ulcer") Follow-up Questions

Provide the duration, in weeks, for each stage of your patient's disease:

t1 \_\_\_\_\_; t2 \_\_\_\_\_; t3 \_\_\_\_\_.

In the tables below, select a value for each cell within the range provided. If you do not select a

value explicitly, the default (in parentheses) will be used. Empty cells are non-applicable.

Among the values you need to select are values for total time in reflux and the corresponding DeMeester score. A crib is provided for general orientation but it is not binding: you may create other associations as well. Note: In general, patients with mild reflux (~ 20%) may be managed with adherence to lifestyle modifications. One would assume that those with reflux less than or equal to 1.5 hrs and a DeMeester score of less than or equal to 20 would fit this profile. This expectation is shown in the table by the contrast between "Total time in reflux with bad habits" and "Total time in reflux with bad habits removed". If your patient has no bad habits, then these should be the same.

DeMeester Crib

Time in Reflux (hours)	FYI: converted to percent time in reflux...	DeMeester Score
0	0	0
.96	4%	7
1.2	5%	10
1.44	6%	18
1.5	6.25%	20
1.68	7%	25
1.92	8%	32
2.4	10%	40
2.88	12%	48
3.0	12.5%	50
3.6	15%	60
4.5	18.75%	70
4.8	20%	80
6.0	25%	120

Physiological Property Values	t1 "inflam."	t2 "erosion"
Total time in reflux	1.2 – 1.44 (1.44)	1.68 – 2.4 (1.68)
DeMeester score	10 – 18 (18)	25 – 40 (32)
Total time in reflux; no bad habits	.96 – 1.44 (.96)	1.68 – 2.4 (1.68)
DeMeester; no bad habits	7 – 18 (7)	25 – 40 (32)
Length of erosion (cm.)*		.5 – 4 (2)
Diam. of erosion (mm.)*		.5 – 3 (1)
Number of erosions*		1 – 4 (2)
Depth of ulcer (mm.)*		
Diameter of ulcer (mm.)*		
Number of ulcers*		

\* The value listed is the end value for the given time period, rather than the start value, as in most tables.

Symptoms	t1 "inflam."	t2 "erosion"
----------	--------------	--------------

Heartburn frequency (#/day)	3 – 5 (4)	6 – 8 (7)	9 – 10 (9)
Heartburn severity	.3 – .5 (.4)	.6 – .8 (.7)	.9 – 1.0 (.9)
Symptom correlation (for pH monitoring)	0 – 1	0 – 1	0 – 1
Regurgitation freq. (#/day)	3 – 5 (4)	6 – 8 (7)	9 – 10 (9)

Treatments (Does it work? y / n)	t1 "inflamm."	t2 "erosion"	t3 "ulcer"
PPI (QD)	y / n	y / n	y / n
PPI (BID)	y / n	y / n	y / n
H2 Blocker	y / n	y / n	y / n
Lifestyle Modifications	y / n	y / n	y / n

### End of survey

As should be clear, VP authors can create patients of very different profiles based on the values they select for each of the properties queried in this survey and its counterparts that cover the other paths of the disease. Naturally, the patient authoring survey does not reflect all of the knowledge encoded in the model that permits real-time simulation of GERD. However, we hope that this snapshot of a second disease provides a glimpse into the generality – and extensibility – of our approach to the modeling of diseases and their treatments.

## 5. Interacting with Virtual Patients

At the beginning of a simulation session, the system presents the user with a virtual patient about whose diagnosis he or she initially has no knowledge. Currently, the user selects diagnostic and treatment options using a menu-based – essentially, multiple-choice – interaction system, though we plan to incorporate natural language-based interaction in the near future (see below for details). In response to user queries, the system returns information stored in the fact repository for the given MVP instance. The system's responses to the user's queries are stored as data in the user's own copy of the MVP. At the beginning of the session, this copy is representative of a virtual human without abnormalities. The diagnosis process results in a gradual modification of the user's copy of the MVP so that in the case of successful diagnosis, it closely resembles the system's version of the MVP. This

format is best described as similar to the game of Battleship, where the players gradually determine the positions of the opponent's ships on a grid. At any point during the diagnosis of the

MVP, the user may proceed with treatment. In other words, the system allows the user not only to issue queries but also to intervene in the simulation, changing property values within the MVP. Any single change can induce other changes – in our terms, trigger a related script. For example, if a user treats an achalasia patient by performing a Heller Myotomy, and if the Heller is complete (as decided earlier by the VP author), then the patient will exit the Heller with GERD caused by wide-open reflux. The chaining of medical events and their effects justifies our modeling of disease biomechanistically since a script for GERD – whether it is caused by an MVP's predispositions or by a complete Heller – will be the same regardless of its triggering biomechanism.

## 6. Tutoring

Although the VP system will ultimately function as a tutor, we began development with the modeling of patients and diseases and are only now beginning to work on the tutoring aspects. With respect to the latter, we have gained useful insights from the CIRCSIM group, whose work we cited in the introduction. CIRCSIM-Tutor's history is actually noteworthy as a point of comparison with the MVP project. The ancestor system for CIRCSIM-Tutor, MacMan, was a mathematical model of the baroreceptor reflex that could be explored by students but provided no feedback. It was found that the lack of feedback made it a non-optimal teaching tool, which led to the development of the first spin-off system, Heartsim, which provided some feedback. However, with the Heartsim system, developers realized that the mathematical model was not being exploited and that the most effective teaching was based on stored correct predictions rather than real-time calculations using the mathematical model. As such, the final system, CIRCIM-Tutor (which is still under development), does not actively use the mathematical model: the



dynamic aspect of the system is constrained to the tutoring process itself. To summarize, this system went from offering students a dynamic mathematical model with no tutoring support to offering them tutoring without the dynamic mathematical model. The contrast with MVP is clear: for us, the autonomous functioning of MVPs is, and will remain, as important as tutoring.

We believe that one of the most valuable aspects of the MVP system will be offering students countless practice cases upon which to hone their diagnostic and treatment skills. In fact, there is evidence that learning by working through computer-based scenarios can be very effective: for example, in the evaluation of the SHERLOCK II system, which teaches electronics troubleshooting, it was reported that technicians learned more from using this system for 24 hours than from 4 years of work in the field (Evens and Michael 2006: 375).

Before launching full-scale work on the MVP project, exploratory observational exercises were conducted with medical students at the University of Maryland Baltimore School of Medicine to understand the specifications for effective interaction with a simulated patient (as reported in Mallott et al. 2005).<sup>1</sup> In the exercises, the students managed several structured patients in electronic and manual simulations. All the exercises employed patient management problems used routinely in teaching and focused on high-level decision-making, such as the proposal and proof of an inference or the substantiation of an intervention. The most notable observations from this and a follow-up study of simulation for medical training were (ibid):

- The simulation must accommodate trial and error patient management with multiple clinically plausible pathways to a solution.
- Changes in patient anatomy and physiology resulting from user action or

disease processes over time must result in a consistent appropriate alteration of the state of the patient.

- The representation of time-related patient activities is critical for successful simulation. This includes allowing the user to “advance the clock” to the next phase of patient management to assess patient responses.

All of these desiderata are being incorporated into the MVP system.

We mentioned earlier that we are planning to incorporate natural language interaction into the system, a desideratum that has recently been expressed by developers of many tutoring systems. Unlike others, however, our group has been working on knowledge-based natural language processing (NLP) for twenty years. In fact, the OntoSem ontology, knowledge representation language, and many of the processors that are serving as a substrate for the MVP system were all originally developed for NLP applications. Therefore, while we would not claim to have the NLP angle tied up (we know the problems too intimately for that), we approach the integration of natural language support as long-time practitioners of this craft.

## References Cited

- Altman, R.B., A. Valencia, S. Miyano, S. Ranganathan. 2001. Challenges for Intelligent Systems in Biology. *IEEE Intelligent Systems* 16(6): 14-20.
- Bodenreider, O. 2004. The Unified Medical Language System (UMLS): Integrating biomedical terminology. *Nucleic Acids Research* 32: 267-270.
- Evens, M. and J. Michael. 2006. *One-on-One Tutoring by Humans and Computers*. New Jersey and London: Lawrence Erlbaum and Associates, Publishers.
- Mallott, D., J. Raczek, C. Skinner, K. Jarrell, M. Shimko and B. Jarrell. 2005. A basis for

---

<sup>1</sup> The exercises were conducted under IRB Exemption No. BJ-090103 for the project entitled “Computer Simulation as an Aid to Enhancing Medical Education.”

electronic cognitive simulation: the heuristic patient. *Surgical Innovations* March, 12(1):43-9.

Nirenburg, S. and V. Raskin. 2004a. *Ontological Semantics*. The MIT Press.

Rosse, C. and J.L.V. Mejino. 2004. A reference ontology for bioinformatics: The Foundational Model of Anatomy. *Journal of Biomedical Informatics*.

## Appendix 3-E

### ACHALASIA

In achalasia, for some unknown reason, the relaxing neurons in the LES die off, which means that the LES is unable to relax sufficiently to allow food to pass through to the stomach. (The “at rest” state of the LES is tight, since it functions to keep stomach contents from spilling out into the esophagus.) The failure of the LES to loosen begins to cause aperistalsis (malfunction of peristalsis) in the lower esophagus. There is a physiological cause-effect relationship between the LES being too tight and the esophageal segments losing their ability to peristalt.

For all achalasia patients, lifestyle events that would typically lower LES pressure (smoking, eating chocolate, etc.) do not affect LES pressure.

Achalasia, like all diseases in the VP system, is described using time-oriented tables that record the physiology, symptoms, expected test results, and treatment effects throughout the course of the disease.

### Basic Physiology of Achalasia

Table 1 represents the changes in physiological properties that take place during achalasia.

- The values in the cells represent starting values for that time period; intermediate values are interpolated using a linear function.
- Physiologically, the independent variable is the ratio of contracting to relaxing neurons in the distal esophagus; all other property values are dependent upon that. However, for purposes of modeling, we refer to the basic LES pressure as a stand-in for the independent variable, since the correspondence between the neuron ratio and basic LES pressure is fixed, and physicians prefer to orient diagnostics and treatment around basic LES pressure.
- The values in red represent parameterizable features which can be set for a given patient by the patient author. The “legal” ranges are presented as guidelines; the defaults, which are used if no values are explicitly selected, are in parentheses.
- Table 1 is used directly as a data source for implementation.

Table 1: Physiological properties that change due to achalasia

	<i>default times in months</i>	<i>t0 (6)</i>	<i>t1 (6)</i>	<i>t2 (6)</i>	<i>t3 (6)</i>	<i>t4 (6)</i>
1.1	Ratio of contracting to relaxing neurons in the distal esophagus	100/100	75/100	50/100	25/100	10/100
1.2	Basic LES pressure (torr)	0 - 40 (25)	$P_{t0} + 10$	$P_{t0} + 20$	$P_{t0} + 30$	$P_{t0} + 40$
1.3	Relaxed LES pressure (torr)	0	10	20	30	40
1.4	Relaxed LES diameter (cm.)	2	1.5	1.0	.5	0
1.5	amplitude of contraction during peristalsis	80	60	35	20	10
1.6	efficacy of peristalsis (reflected	<i>normal</i>	<i>normal</i>	<i>intermittent-</i>	<i>aperistalsis</i>	<i>aperistalsis</i>

	in changes during swallowing)			<i>peristalsis</i>		
1.7	diameter of distal esophagus (cm.)	2	3	4	5	6
1.8	bolus quantity in distal esophagus (on abstract scale 0-1)	0	0 - .2	.2 - .4	.4 - .7	> .7
1.9	duration of swallowing (min.)	0	1	5	10	35000

## Symptom Profile of Achalasia

The basic symptom profile for achalasia is recorded in Table 2.

- The durations of  $t_0 - t_1$  are copied from Table 1.
- Since the experiencing of symptoms varies widely across patients and cannot be directly linked to given physiological states, the symptom profile for each patient is typically overtly specified by the patient author. However, defaults are provided as well.
- The values in the cells represent starting values for that time period; intermediate values are interpolated using a linear function.
- The symptoms table is used directly as a data source during implementation.

Table 2: Symptom profile of achalasia

		t0	t1	t2	t3	t4
2.1	Difficulty swallowing	0	0 - .2 (.1)	.1 - .4 (.2)	.2 - .7 (.6)	.5 - 1 (.9)
2.2	Weight loss	0	0 - .05 (0)	0 - .1 (0)	0 - .15 (.1)	.05 - .2 (.2)
2.3	Chest pain	0	0 - .3 (.1)	0 - .5 (.3)	.3 - .8 (.5)	.5 - 1 (.7)
2.4	Regurgitation (times/month)	0	0 - 4 (0)	0 - 20 (10)	20 - 50 (40)	20 - 100 (70)

## Test Results

We elicit relevant tests and test results in tabular form only to facilitate knowledge elicitation. The test results table (Table 3) does not serve as a formal specification nor is it used as a source of data for implementation because test results (unlike symptoms) are completely predictable based on physiological properties of the VP. As such, they are not asserted by the patient author, they are automatically generated by the simulator.

Table 3. Coarse-grained inventory of expected test results during the stages of achalasia; for use by SMEs and knowledge engineers only.

		t0	t1	t2	t3	t4
3.1	EGD	<i>normal</i>	Dilated esophagus.	Dilated esophagus.	Dilated esophagus. Narrowing and pop upon entering.	Dilated esophagus. Narrowing and pop upon entering.

					Retained debris in esophagus.	Retained debris in esophagus.
3.2	Manometry	Incomplete relaxation of LES.	Incomplete relaxation of LES.  High-normal LES pressure.	Incomplete relaxation of LES.  High-normal or hypertensive LES.  Intermittent peristalsis.	Incomplete relaxation of LES.  Hypertensive LES.  Aperistalsis	Incomplete relaxation of LES.  Hypertensive LES.  Aperistalsis
3.3	Barium swallow	<i>normal</i>	Delayed emptying.	Delayed emptying.	Bird's beak.  Dilated esophagus.  Retained debris.  Retained barium.	Bird's beak.  Dilated esophagus.  Retained debris.  Retained barium.

Once this table has been compiled, subject-matter experts (SMEs) must indicate which physiological properties give rise to each result so that, when a test is ordered, the simulator can check through the inventory of possible results and return only those that apply to the given VP at the given time.

The physiological prerequisites for each test result are recorded in the **Test Results Repository (TRR)**, which is a knowledge base whose structure and relationship to the ontology are similar to that of a lexicon:

- head words/phrases are strings (i.e., components of natural language, not ontology)
- the meaning of these strings is described using ontological concepts, properties and values.

The TRR, like a lexicon, acts as a bridge between expressions in natural language and expressions in the ontological metalanguage. For example, one result of an esophagogastroduodenoscopy (EGD) is “dilated esophagus”, which is a natural language description of the VP state when the diameter of his distal esophagus is greater than 3 centimeters. Although physicians need to be able to refer to a dilated esophagus, it would be incorrect to make this into an ontological concept because it precisely matches the compositional ontological representation

(DIAMETER  
(DOMAIN (value LOWER-BODY-OF-ESOPHAGUS))  
(RANGE (> 3 (default-measure CENTIMETER))))

In the TRR, all test results are indexed as strings and are mapped to the ontological concept TEST-RESULT-AS-ENTITY, which is a child of REPRESENTATIONAL-OBJECT (note that TEST-RESULT is already used as the name of a property). The inventory of property values that must be defined in the TRR for each test result are:

- DEFINITION (for human use only)
- TEST-RESULT-OF, which links the test result to the test that can give rise to it
- SPECIALISTS-INTERPRETATION, whose filler is one or more text strings that would be returned by a diagnostician reporting the given result; if more than one is provided, one is selected at random each time the given result is returned.
- TEST-RESULT-CONTENT, which describes, using the ontological metalanguage, the precondition(s) for returning the test result.

Below are two sample entries from the TRR.

(dilated-esophagus

(definition "an abnormally wide lower esophagus, defined as a diameter greater than 3 cm.")

(TEST-RESULT-AS-ENTITY

(SPECIALISTS-INTERPRETATION "dilated esophagus")

(TEST-RESULT-OF ESOPHAGOGASTRODUODENOSCOPY)

(TEST-RESULT-CONTENT

(DIAMETER

(DOMAIN (value LOWER-BODY-OF-ESOPHAGUS))

(RANGE (> 3 (default-measure CENTIMETER))))))

(hypertensive-les

(definition "a condition in which the basic LES pressure is abnormally high: > 35 torr")

(TEST-RESULT-AS-ENTITY

(SPECIALISTS-INTERPRETATION "hypertensive LES")

(TEST-RESULT-OF ESOPHAGEAL-MANOMETRY)

(TEST-RESULT-CONTENT

(PRESSURE

(DOMAIN (value LES)

(RANGE (> 35 (measured-in torr))))))

Just as lexical mappings for different languages are shown in the ontology, so are head words/phrases from the TLL. For example, Figure 1 shows the current inventory of results for ESOPHAGOGASTRODUODENOSCOPY (i.e., EGD) (these actually should be in lower case since they are not concept names).

Concept: ESOPHAGOGASTRODUODENOSCOPY		
DEFINITION	VALUE	examination of the lumen of the esophagus, stomach and duodenum using an endoscope (also called EGD)
IS-A	VALUE	ENDOSCOPY
AGENT	SEM	GASTROENTEROLOGIST
INSTRUMENT	DEFAULT	UPPER-ENDOSCOPE
CREATES-IMAGE-OF	SEM	CAVITY-OF-STOMACH
		LUMEN-OF-DUODENUM
		LUMEN-OF-ESOPHAGUS
		DILATED-ESOPHAGUS
MEDICAL-TEST-RESULT-STRING	SEM	EGD-NORMAL
		EROSION-IN-ESOPHAGUS
		ESOPHAGEAL-INFLAMMATION
		NARROWING-OF-LES-WITH-POP-UPON-ENTERING-STOMACH
		PEPTIC-STRICTURE-TEST-RESULT
		RETAINED-DEBRIS-IN-ESOPHAGUS
		SMOOTH-NARROWING-OF-DISTAL-ESOPHAGUS
		TUMOR-IN-ESOPHAGUS
		ULCER-IN-ESOPHAGUS

Table 1.

Let us reiterate how test results are handled in simulation. Whenever a student orders a diagnostic procedure, the simulator scans the TRR for test results derivable from that procedure. It then checks to see for which one(s) the VP has the requisite property values and, for each applicable result, returns one of the text strings listed in the field SPECIALISTS-OPINION. In this way, test results automatically derive from the physiology of the VP. So, while it is useful for SMEs and knowledge engineers to create a test results table as a crib for keeping track of diagnostics and their expected results, that table-based information is not consulted by the simulator, whose only interest is the property values that define the VP at a given time.

## Treatments

Treatments can be categorized by various parameters, including:

- **dependency or lack thereof on the stage of the disease:** some treatments have different outcomes depending on the patient's disease stage, while others can affect patients similarly no matter what the disease stage
- **duration of effect:** some treatments are always definitive (e.g., removing a spleen), some always wear off (e.g., BoTox) and some may or may not be definitive (e.g., Heller Myotomy).

Different kinds of treatments require slightly different knowledge representation formats.

An important aspect of modeling treatments is that different responses to treatment can be modeled by changing values for only a few key variables. For example, in achalasia, the basic LES pressure is the anchor variable – the stand-in for the actual independent variable, which is the ratio of relaxing to contracting neurons (a property that physician's don't refer to this in describing, diagnosing or treating achalasia). After any treatment, the value for the

basic LES pressure must be asserted by the VP author (or else the default can be used). Most other physiological property values, and all symptoms, remain in a constant relationship with basic LES pressure. As such, if a VP has a basic LES pressure of 25 early in his disease, and again late in his disease after some treatment has been administered, most of his other physiological property values and all of his symptoms will be the same at both times. Therefore, the original disease and symptoms tables for the given VP continue to be used as look-up tables even after a patient receives a treatment. Referring to the original tables is not only a convenient shorthand, it ensures consistency of the patient profile throughout the patient's existence (it would be unrealistic for him to have vastly different symptom profiles at different times given the same underlying physiological state).

## BoTox

Botox reduces the number of contracting neurons in the LES, so the ratio of contracting neurons to relaxing neurons improves. (Actually, BoTox poisons all neurons, but the net effect is still to improve the ratio.) If BoTox is effective (it is not effective for all patients at all disease stages) its effects typically last for 6-12 months. In the current implementation, the durations are exactly 0, 6 or 12 months, but other durations could be included as well if deemed clinically and pedagogically relevant.

BoTox does not cure achalasia and does not stop disease progression. Therefore, when its effects wear off, the disease will be at the same stage as it would have been if BoTox had not been administered.

The BoTox table shows the typical **initial effects** of BoTox on patients experiencing different stages of achalasia. The table only shows initial effects after treatment and duration of treatment; everything else about disease progression after BoTox is interpolated based on the patient's original disease profile. Therefore, the table should not be read left-to-right, as implied by the wide black lines separating columns.

The simulator can calculate what the disease state would be at any point during the treatment based on two types of information: property values at BoTox-Begin (in the table) and property values at BoTox-End (the point in the disease where the VP would have been anyway if no BoTox had been given). Values for BoTox-End are drawn from the original disease table. A linear change in values from BoTox-Begin to BoTox-End is assumed. This calculated change in the VPs state *is* the treatment script – which is slightly different for each VP.

The four properties explicitly listed for each patient after BoTox treatment are listed because they cannot be inferred from the basic disease table. Obviously, the basic LES pressure must be asserted, but the relaxed LES pressure must also be asserted because the ratio between basic and relaxed is not necessarily the same once the person has achalasia as it was before he was diseased (even if the basic LES pressure goes down to what looks like a normal range).

Similarly, we need to assert the esophageal contents *or* emptying delay *or* both. The relationship between esophageal contents and emptying delay is sufficiently fixed so that one can be derived from the other using the original disease table as a guide. If both are asserted, then both are used without reference to the original correlation between them.

If a patient is given BoTox during this stage of the disease,	t0 (vals. not	t1	t2	t3	t4
---	------------------	----	----	----	----



	confirmed)				
his relaxed LES pressure will initially go down to	0 – 5 ()	3 – 10 ()	5 – 12 ()	18 – 22 ()	28-35 ()
his basic LES pressure will initially go down to	10 – 30 ()	15 – 40 ()	25 – 50 ()	35 – 60 ()	45-65 ()
the effect will wear off over # months	0, 6, 12 ()	0, 6, 12 ()	0, 6, 12 ()	0, 6, 12 ()	0, 6, 12 ()
his esophageal contents will initially go down to	0 ()	0 - .15 ()	0 - .2 ()	.2 - .4 ()	.5 - .6 ()

## Pneumatic Dilation

Pneumatic Dilation (PD) is an endoscopic procedure during which a balloon is inflated in the LES to rip the muscle layer. PD tends not to reduce basic LES to 0 (as a Heller Myotomy sometimes can); a resulting basic pressure of 10-12 is considered a successful result. After PD, scar tissue forms at some rate but there is no data on which to base a specific model of scar tissue formation so this is not included in the current VP model.

There are 3 clinically relevant scenarios representing the efficacy of PD. All are possible regardless of the disease stage at which PD is carried out:

1. definitive cure (great success that lasts long term);
2. success but regression;
3. failure.

Moderate success is not an interesting outcome and is not necessary for teaching purposes. If the results of PD are only moderate, one repeats the PD right away or follows up with a Heller Myotomy until a good result is achieved.

We model the results of PD based on clinical knowledge of patient symptoms over time since no explicit tests are done at the time of PD to see how well it has worked.

The time frames of interest for post-PD patients are preset to the clinically relevant intervals of 1 month, 1 year and 5 years.

If the PD fails for a given patient, all of his property values stay as they were, as does the state of his disease.

If the PD is at least initially successful (scenario 1 or 2), certain property values of the VP need to be changed either explicitly by the VP author or according to the default values.

### Scenario 1, parameterizable property values

	after dilation
Basic LES pressure	10 – 12 (10)
Relaxed LES pressure	0 – 1.5 (.5)
Esophageal Contents	0 - .2 (.1)
Emptying Delay (mins.)	0 – 5 (1)

### Scenario 2, parameterizable property values

	time of PD	PD + 1 month	PD + 1 year	+ 5 years
Basic LES	10 – 12 (10)	11 – 15 (15)	30 – 40 (30)	55 – 75 (55)
Relaxed LES	0 – 5 (3)	5 – 10 (8)	20 – 30 (20)	40
Esophageal contents	0 - .2 (.1)	0 - .2 (.2)	.6	> .7
Emptying delay (mins)	0 – 5 (1)	0 – 5 (3)	10	35000

## Heller Myotomy

Heller Myotomy is surgery that cuts the LES more neatly (and therefore with less subsequent scar tissue) than PD. In theory, the muscle should be cut 100%, leaving a basic LES pressure of 0. Such a result is called a complete Heller and, by definition, it will make the patient a GERD patient with wide open reflux (there are anti-reflux procedures that can be done at the time of HM but we're not including them yet). Incomplete Heller Myotomies leave some muscle tissue, which then can regenerate. In 70% of patients the results of a HM last for 10 years.

For the majority of achalasia patients from 30-60 years old, laparoscopic Heller Myotomy is considered the best care, given there are good surgeons available.

As with PD, there are 3 clinically relevant scenarios representing the efficacy of HM. All are possible regardless of the disease stage at which HM is carried out:

1. definitive cure (great success that lasts long term);
2. success but regression;
3. failure.

**Heller Myotomy Scenario 1, parameterizable property values**

<i>HM Scenario 1 Table</i>	<b>after dilation</b>
Basic LES pressure	0 – 10 (3)
Relaxed LES pressure	0 – 5 (1)
Esophageal Contents	0 - .2 (.1)
Emptying Delay (mins.)	0 – 5 (1)

**Heller Myotomy Scenario 2, parameterizable property values**

<i>HM Scenario 2 Table</i>	time of HM	HM + 1 month	HM + 1 year	+ 5 years
Basic LES	0 – 10 (3)	2 – 15 (5)	25 – 40 (25)	40 – 65 (45)
Relaxed LES	0 – 5 (1)	5 – 10 (5)	15 – 25 (15)	40
Esophageal contents	0 - .2 (.1)	0 - .2 (.1)	.3 - .6 (.3)	> .7
Emptying delay (mins)	0 – 5 (1)	0 – 5 (2)	6 – 10 (7)	35000

# Achalasia Patient Authoring Questionnaire

<i>Basic Data Table</i>	
Last name	
First name	
Middle name	
Age	
Gender	
Race	
Basic LES pressure at t0	
t0 duration (weeks)	
t1 duration (weeks)	
t2 duration (weeks)	
t3 duration (weeks) *	
T-time when patient presents	
Description of this patient and its teaching goals (any length; just for people, not to be processed)	

\* t4 is understood to start at the end of t3 and last forever after.

In all the tables below, the ranges of values in white cells act as a guide. In the orange cell below each white cell, fill in an actual value (not a range) for your patient at the *start* of the given time period. Interpolation of changes throughout a time period, based on the start value of the next time period, will be done by the simulator. If you do not fill in a value explicitly, the default value will be used (written in parentheses next to the range). Hit the Tab key to jump to next cell in given row.

## Parameterizable Physiological Properties

<i>Symptoms Table</i>	<b>t0</b>	<b>t1</b>	<b>t2</b>	<b>t3</b>	<b>t4</b>
<b>Esophageal Contents</b>	0	0 - .2 (.1)	.2 - .4 (.3)	.5 - .7 (.6)	> .7 (.8)
<i>Esoph. cont. actual</i>					

## Symptoms

<i>Symptoms Table</i>	<b>t0</b>	<b>t1</b>	<b>t2</b>	<b>t3</b>	<b>t4</b>
<b>Difficulty swallowing</b>	0	0 - .2 (.1)	.1 - .4 (.2)	.2 - .7 (.6)	.5 - 1 (.9)
<i>Dif. swal., actual</i>					
<b>Weight loss</b>	0	0 - .05 (0)	0 - .1 (0)	0 - .15 (.1)	.05 - .2 (.2)
<i>Wt. loss, actual</i>					
<b>Pain in chest</b>	0	0 - .3 (.1)	0 - .5 (.3)	.3 - .8 (.5)	.5 - 1 (.7)
<i>Pain in chest, actual</i>					
<b>Regurgitation (#/mo.)</b>	0	0 - 4 (0)	0 - 20 (10)	20 - 50 (40)	20 - 100 (70)
<i>Regurgitation, actual</i>					

### Treatment with Botox (no defaults)

The effect of Botox is calculated by using the starting values asserted here, and interpolating the rest of the patient's values over time based on his/her basic rate of disease progression, described earlier (BoTox does not affect that). Therefore, each column of this table should be understood as a different "treatment path": read only up-to-down, not left-to-right (thus, the thick black lines between columns).

<i>Botox Table</i>	<b>t1</b>	<b>t2</b>	<b>t3</b>	<b>t4</b>
<b>Relaxed LES pressure</b>	0 – 5	5 – 12	18 – 22	28-35
<i>Relaxed LES, actual</i>				
<b>Basic LES</b>	15 – 40	25 – 50	35 - 60	45-65
<i>Basic LES, actual</i>				
<b>Duration of Effect (mos.)</b>	0, 6, 12	0, 6, 12	0, 6, 12	0, 6, 12
<i>Dur. of eff., actual</i>				
<b>Esophageal contents</b>	0 - .15	0 - .2	.2 - .4	.5 - .6
<i>Esoph. cont., actual</i>				

### Treatment with Pneumatic Dilation:

There are three basic scenarios for how a patient might respond to pneumatic dilation. Select **only one** for your patient. If you select scenarios 1 or 2, answer the follow-up questions **only for that scenario**.

<i>PD Scenario Table</i>	<i>Type "yes" for the selected one</i>
1. Effective forever, no regression	
2. Effective but regression over time	
3. Ineffective	

### Scenario 1, follow-up questions

<i>PD Scenario 1 Table</i>	<b>after dilation</b>
<b>Basic LES pressure</b>	10 – 12 (10)
<i>basic LES pressure, actual</i>	
<b>Relaxed LES pressure</b>	0 – 1.5 (.5)
<i>relaxed LES pressure, actual</i>	
<b>Esophageal Contents</b>	0 - .2 (.1)
<i>esoph. contents, actual</i>	
<b>Emptying Delay (mins.)</b>	0 – 5 (1)
<i>emptying delay., actual</i>	

### Scenario 2, follow-up questions

Note: unlike with Botox, this treatment fundamentally changes the patient so that the old disease script is no longer applicable. Fill in all cells as a timed progression for your patient.

<i>PD Scenario 2 Table</i>	time of PD	PD + 1 month	PD + 1 year	+ 5 years
<b>Basic LES</b>	10 – 12 (10)	11 – 15 (15)	30 – 40 (30)	55 – 75 (55)
<i>Basic LES, actual</i>				
<b>Relaxed LES</b>	0 – 5 (3)	5 – 10 (8)	20 – 30 (20)	40
<i>Relaxed LES, actual</i>				
<b>Esophageal contents</b>	0 - .2 (.1)	0 - .2 (.2)	.6	> .7
<i>Esoph. cont., actual</i>				
<b>Emptying delay (mins)</b>	0 – 5 (1)	0 – 5 (3)	10	never
<i>empt. delay, actual</i>				

### Treatment with Heller Myotomy:

There are three basic scenarios for how a patient might respond to Heller Myotomy. Select **only one** for your patient. If you select scenarios 1 or 2, answer the follow-up questions **only for that scenario**.

<i>Heller Myotomy Scenario Table</i>	<i>Type “yes” for the selected one</i>
1. Effective forever, no regression	
2. Effective but regression over time	
3. Ineffective	

### Heller Myotomy Scenario 1, follow-up questions

<i>HM Scenario 1 Table</i>	<b>after dilation</b>
<b>Basic LES pressure</b>	0 – 10 (3)
<i>basic LES pressure, actual</i>	
<b>Relaxed LES pressure</b>	0 – 5 (1)
<i>relaxed LES pressure, actual</i>	
<b>Esophageal Contents</b>	0 - .2 (.1)
<i>esoph. contents, actual</i>	
<b>Emptying Delay (mins.)</b>	0 – 5 (1)
<i>emptying delay., actual</i>	



### Heller Myotomy Scenario 2, follow-up questions

Note: unlike with Botox, this treatment fundamentally changes the patient so that the old disease script is no longer applicable. Fill in all cells as a timed progression for your patient.

<i>HM Scenario 2 Table</i>	time of HM	HM + 1 month	HM + 1 year	+ 5 years
<b>Basic LES</b>	0 – 10 (3)	2 – 15 (5)	25 – 40 (25)	40 – 65 (45)
<i>Basic LES, actual</i>				
<b>Relaxed LES</b>	0 – 5 (1)	5 – 10 (5)	15 – 25 (15)	40
<i>Relaxed LES, actual</i>				
<b>Esophageal contents</b>	0 - .2 (.1)	0 - .2 (.1)	.3 - .6 (.3)	> .7
<i>Esoph. cont., actual</i>				
<b>Emptying delay (mins)</b>	0 – 5 (1)	0 – 5 (2)	6 – 10 (7)	never
<i>empt. delay, actual</i>				

# Actual Patient Instance: Implemented and Tested

<i>Basic Data Table</i>	
Last name, First name	Chu, Charles
Age, Gender, Race	40, m, Asian
Basic LES pressure at t0	16
t0, t1, t2, t3 (in months)	8, 6, 6, 6
Description of this patient and its teaching goals (any length; just for people, not to be processed)	fast progressing; many high-extreme values; botox ineffective; pneumatic dilation and Heller Myotomy both have regression; HM not very “complete” (LES down to 10)

<i>Parameterizable Physio Table</i>	<b>t0</b>	<b>t1</b>	<b>t2</b>	<b>t3</b>	<b>t4</b>
Esophageal contents	0	.2	.4	.6	.9

<i>Symptoms Table</i>	<b>t0</b>	<b>t1</b>	<b>t2</b>	<b>t3</b>	<b>t4</b>
Difficulty swallowing	0	.2	.4	.7	1
Weight loss	0	.05	.1	.15	.2
Pain in chest	0	.3	.5	.7	1
Pain on swallowing	0	.3	.5	.7	1
Regurgitation (#/mo.)	0	4	18	48	98

**Treatment with Botox:** ineffective

**Treatment with Pneumatic Dilation:** Effective but regression over time

<i>PD Scenario 2 Table</i>	time of PD	PD + 1 month	PD + 1 year	+ 5 years
Basic LES	12	15	38	65
Relaxed LES	5	10	28	40
Esophageal contents	.2	.2	.6	.9
Emptying delay (mins)	3	5	10	never

**Treatment with Heller Myotomy:** Effective but regression over time

<i>PD Scenario 2 Table</i>	time of PD	PD + 1 month	PD + 1 year	+ 5 years
----------------------------	------------	--------------	-------------	-----------

Basic LES	10	15	38	65
Relaxed LES	5	10	25	40
Esophageal contents	.2	.2	.6	.9
Emptying delay (mins)	3	5	10	never

## Appendix 3-F

# GERD

1. What is GERD
2. How we're not modeling GERD
3. How we are modeling GERD
  - 3.1 Inflammation-erosion-ulcer-peptic stricture
  - 3.2 Inflammation-Barrett's Esophagus-Tumor
  - 3.3 Episodic instances of reflux (as opposed to the long-term disease GERD)
  - 3.4 GERD Test Results Repository

## 1. What is GERD?

GERD is a disease involving irritation of the mucosal lining of the esophagus that results from repeated occurrences reflux. It can take many paths based on an individual's predispositions.

## 2. How we're not modeling GERD

It is not realistic to attempt to model GERD by tracking every GERD-affecting lifestyle activity over weeks and months – or even over 24 hours. We don't have enough concrete, minute-by-minute evidence to indicate how much acid is refluxed with each intake of “irritating” food, what effect each instance of lying down with a full stomach has, how long refluxed acid stays in the esophagus before getting washed down by swallowing, etc. There's no reasonable way to “make it all up” so that the numbers consistently pan out right. So we're not taking a purely physiological cause-effect approach.

## Appendix 3-F

Another modeling idea that turned out to be too complex in its details was modeling individual instances of reflux using a formula that would compare LES pressure with gastric pressure under many different conditions (eating GERD-irritating food, lying down, etc.). The problem is the combinatorics: if caffeine and mints and smoking all cause a lessening of LES pressure, then what is their combined effect? What if the person additionally lies down with a full stomach after eating all of that? The formulas we were constructing were not fine-grained enough to make sensible outcomes regardless of the data combinations.

We had thought of adding patient-specific values for how often and how intensely the given patient experiences pain (from inflammation or pH) when the given nerve receptors are fired. However, this was decided to be not medically useful (nobody really understands how this works among patients), so all patients will experience pain 15% of the time that receptors are fired.

### 3. How we are modeling GERD

There are three meta-scenarios for GERD progression:

1. inflammation – erosion – ulcer – peptic stricture
2. inflammation – Barrett’s esophagus – tumor
3. proximal GERD *{{ this one is postponed till after the demo }}*

From these three meta-scenarios we derive 7 actual scenarios, since different patients have different predispositions for disease progression. That is, some patients never progress past the stage of “inflammation”, some get an erosion but it never turns into an ulcer, some have Barrett’s esophagus but never get a tumor, etc.

The actual scenarios are recorded in the following table, along with the physiological property values of a patient experiencing the given scenario. The description of the scenarios (column 2) indicate the progression and end point of the disease: e.g., inflammation means “just inflammation – no further disease possible in this patient no matter how long he has GERD”.

Basic GERD Scenario	Scenarios	Properties	Values
---------------------	-----------	------------	--------

### Appendix 3-F

Table			
1	inflammation	predisposition-to-erosion predisposition-to-ulcer predisposition-to-stricture predisposition-to-Barretts predisposition-to-tumor predisposition-to-proximal-reflux	no no no no no no
2	inflammation erosion	predisposition-to-erosion predisposition-to-ulcer predisposition-to-stricture predisposition-to-Barretts predisposition-to-tumor predisposition-to-proximal-reflux	yes no no no no no
3	inflammation erosion ulcer	predisposition-to-erosion predisposition-to-ulcer predisposition-to-stricture predisposition-to-Barretts predisposition-to-tumor predisposition-to-proximal-reflux	yes yes no no no no
4	inflammation erosion ulcer peptic-stricture	predisposition-to-erosion predisposition-to-ulcer predisposition-to-stricture predisposition-to-Barretts predisposition-to-tumor predisposition-to-proximal-reflux	yes yes yes no no no
5	inflammation Barrett's	predisposition-to-erosion predisposition-to-ulcer predisposition-to-stricture predisposition-to-Barretts predisposition-to-tumor	no no no yes no

## Appendix 3-F

		predisposition-to-proximal-reflux	no
6	inflammation Barrett's tumor	predisposition-to-erosion predisposition-to-ulcer predisposition-to-stricture predisposition-to-Barretts predisposition-to-tumor predisposition-to-proximal-reflux	no no no yes yes no
7	proximal_reflux	predisposition-to-erosion predisposition-to-ulcer predisposition-to-stricture predisposition-to-Barretts predisposition-to-tumor predisposition-to-proximal-reflux	no no no no no yes

When creating a VP instance, the VP author chooses a scenario and then fills in the applicable follow-up information about his patient at each stage of the scenario.

For this exposition, we discuss the two scenarios that represent the most advanced cases of the two disease paths currently implemented: inflammation-erosion-ulcer-peptic stricture, and inflammation-Barrett's-tumor. The scenarios that represent subsets of these are treated similarly.

### 3.1 Inflammation-erosion-ulcer-peptic stricture (or, for short “To Stricture”)

#### The “To Stricture” physiological table

The physiological table presents the stages of the disease along the x-axis and the relevant properties along the y-axis. The duration for each stage for each patient is set by the VP author.

Physiological Property Values	t1 “inflam.”	t2 “erosion”	t3 “ulcer”	t4 “stricture”
Total time in reflux (hrs.)	1.2 – 1.44 (1.44)	1.68 – 2.4 (1.92)	2.4 – 3.6 (2.88)	1.68 – 3.6 (3.6)

### Appendix 3-F

DeMeester score	10 – 18 (18)	25 – 40 (32)	40 – 60 (48)	25 – 60 (60)
Total time in reflux; no bad habits (hrs.)	.96 – 1.44 (.96)	1.68 – 2.4 (1.92)	2.4 – 3.6 (2.88)	1.68 – 3.6 (3.6)
DeMeester; no bad habits	7 – 18 (7)	25 – 40 (32)	40 – 60 (48)	25 – 60 (60)
Length of erosion (cm.); [ <i>end value</i> ]		.5 – 4 (2)		
Diam. of erosion (mm.); [ <i>end value</i> ]		.5 - 3 (1)		
Number of erosions; [ <i>end value</i> ]		1 – 4 (2)		
Depth of ulcer (mm.); [ <i>end value</i> ]			1 – 3 (2)	
Diameter of ulcer (mm.); [ <i>end value</i> ]			4-10 (5)	
Number of ulcers; [ <i>end value</i> ]			1 – 4 (2)	
Diam. of T10 lumen (cm.); [ <i>end value</i> ]				.5

The most important physiological property of the GERD patient is his total time in reflux, listed first in the table. As always, ranges for values are provided for each stage of the disease, from which the VP author selects actual values for each patient.

For the first 4 properties, the selected values represent the starting values for each stage, with values during the stage being interpolated. Unless otherwise specified, indicated values in our disease tables are always starting values.

For the next 6 properties, by contrast, the selected values represent the ending values for each stage (which is why [*end value*] is indicated). Why? Because the “erosion”, “ulcer” and “stricture” stages instantiate new objects and those objects, by their very nature, start out extremely small then grow to some maximum size for the given patient (we are modeling them as starting at .001, with the scale indicated – cm. or mm.).

The final property value – the diameter of the lumen of the T10 segment of the esophagus – also represents the ending value, since the lumen starts out at its regular size (2 cm.) and is made narrower by the stricture.

The next property is the patient’s DeMeester Score, which is a composite score that would be returned if pH Monitoring were carried out on the patient. Unlike most test scores, which are automatically derived from property values of the VP, this test score needs to be asserted by the VP author because it depends upon a complex calculation of many physiological factors. Since there is no pedagogical utility in modeling that, physicians assert the expected scores based on clinical knowledge. There is some correlation between



## Appendix 3-F

DeMeester score and total time in reflux. As a guide for VP authors, we present a non-binding crib representing these correspondences.

DeMeester Crib

Time in Reflux (hours)	FYI: converted to percent time in reflux...	DeMeester Score
0	0	0
.96	4%	7
1.2	5%	10
1.44	6%	18
1.5	6.25%	20
1.68	7%	25
1.92	8%	32
2.4	10%	40
2.88	12%	48
3.0	12.5%	50
3.6	15%	60
4.5	18.75%	70
4.8	20%	80
6.0	25%	120

The next two properties are total time in reflux and DeMeester score if the patient adheres to lifestyle modifications. For example, if a heavy coffee drinker stops drinking coffee, his GERD symptoms might disappear. In general, some patients with mild reflux (total time in reflux of less than or equal to 1.5 hours and a DeMeester score of less than or equal to 20) can be managed with adherence to lifestyle modifications. If a patient does not have any GERD-irritating lifestyle habits, or if stopping those habits has no effect on his GERD, then the values for these pairs of properties (with and without bad habits) can be the same.

When the inflammation stage is over, inflammation – which is a property value of the mucosa – remains and an erosion object is instantiated, with a cardinality of 1 and all of its size measures set to .001. How many erosions the VP will experience and their

## Appendix 3-F

maximum sizes along several dimensions is set by the VP author and recorded in the table. These maximum values will be reached (if, of course, the patient is not successfully treated for GERD in the interim) at the end of the erosion stage, at which point the erosion object(s) are destroyed and corresponding ulcer object(s) are created, whose size also starts at .001. The cardinality of ulcers is fixed to be the same as the cardinality of erosions (a simplification that does not impede learning goals). When the ulcer stage is over, all ulcer objects are destroyed and a stricture object is created. The stricture occupies space in the lumen, eventually reducing it to a size of .5 for all patients if they are not treated beforehand.

### The “To Stricture” Symptoms Table

Symptoms for GERD patients, like their physiology, derives from clinical knowledge. As noted earlier, patients do not experience symptoms every time they have reflux, so there is no direct correlation between physiology and symptoms. The symptoms table, therefore, records general patterns of reported symptoms.

Symptoms	t1 “inflam.”	t2 “erosion”	t3 “ulcer”	t4 “stricture”
Heartburn frequency (#/day)	3 – 5 (4)	6 – 8 (7)	9 – 10 (9)	6 – 10 (7)
Heartburn severity	.3 – .5 (.4)	.6 – .8 (.7)	.9 – 1.0 (.9)	.6 – 1.0 (.7)
Symptom correlation (for ph monitoring)	0 – 1	0 – 1	0 – 1	0 – 1
Regurgitation freq. (#/day)	3 – 5 (4)	6 – 8 (7)	9 – 10 (9)	6 – 10 (7)

The only property that requires further description is “symptom correlation”, which refers to how closely the patient’s symptoms match actual cases of reflux as recorded during pH monitoring. A correlation of greater than 50% is considered normal.

### The “To Stricture” Treatments Table

The treatments for GERD are listed in the table below. Any of the first four (lifestyle modifications or medication) can be administered to patients at any time, and might be effective or ineffective. We use y/n (“Is it effective?”) in the table rather than the more opaque e/i. The last treatment effect represents the final diameter of the LES if the patient undergoes TTS dilation to correct the effects of a peptic stricture. This procedure can be more effective (the LES diameter is wider) or less effective (it is narrower).

Treatments (does it work? y / n)	t1 “inflam.”	t2 “erosion”	t3 “ulcer”	t4 “stricture”
Lifestyle Modifications	y / n	y / n	y / n	y / n

## Appendix 3-F

H2 Blocker	y / n	y / n	y / n	y / n
PPI (QD)	y / n	y / n	y / n	y / n
PPI (BID)	y / n	y / n	y / n	y / n
LES diameter after TTS dilation				1 - 3

When a successful treatment is administered, the path of GERD is reversed, and the disease is cured at a faster rate than it originally developed. That is, even if it took a patient 2 years to develop an ulcer, with symptoms worsening gradually over time, an effective treatment will cure him in a few weeks' time:

- first, over a period of 12 weeks, his ulcer(s) will disappear and will be replaced by new erosion object(s);
- next, over a period of 8 weeks, the erosion(s) will diminish in size until they disappear, leaving only inflammation;
- finally, over a period of 2 weeks, the inflammation will also stop, changing the value of the “inflamed” property of the mucosa from “yes” to “no”.

A stricture cannot be reversed except by TTS dilation. So, while a patient's GERD symptoms (heartburn, etc.) can be treated using medication, only a TTS will change the diameter of the esophagus that is the site of stricture.

### 3.2 Inflammation-Barrett's Esophagus-Tumor (or, for short “To Tumor”)

The tables for the “to tumor” scenarios largely parallel those for the “to stricture” scenarios. The main difference is that, thus far, we have not included how tumors are treated or their further implications – once a tumor is detected, the student stops work. If a patient is treated for GERD when he already has Barrett's or tumor, those conditions do not go away due to the GERD treatment; the GERD treatment only treats the effects of inflaming the mucosa of the esophagus.

Basic physio table “to tumor”	t1 “inflam.”	t2 “Barrett's”	t3 “tumor”
Total time in reflux (hrs.)	1.2 – 1.44 (1.44)	1.68 – 2.4 (1.92)	1.68 – 2.4 (1.92)
DeMeester score	10 – 18 (18)	25 – 40 (32)	25 – 40 (32)
Total time in reflux; no bad habits	.96 – 1.44 (.96)	1.68 – 2.4 (1.92)	1.68 – 2.4 (1.92)
DeMeester; no bad habits	7 – 18 (7)	25 – 40 (32)	25 – 40 (32)

## Appendix 3-F

Size of residual lumen (cm.) [ <i>end value</i> ]			.5
---	--	--	----

Symptoms Table “to tumor”	t1 “inflam.”	t2 “Barrett’s”	t3 “Tumor”
Dysphagia	0 - .3 (.1)	0 - .3 (.1)	<i>calculated</i>
Heartburn frequency (#/day)	3 – 5 (4)	0 - 8 (4)	0 - 8 (4)
Heartburn severity	.3 – .5 (.4)	.1 - .8 (.4)	.1 - .8 (.4)
Symptom correlation (for ph monitoring)	0 - 1	0 – 1	0 - 1
Regurgitation frequency (#/day)	3 – 5 (4)	6 – 8 (7)	6 – 8 (7)

Treatment “to tumor”	t1 “inflam.”	t2 “Barrett’s”	t3 “tumor”
PPI QD	y / n	y / n	y / n
PPI BID	y / n	y / n	y / n
H2 Blocker	y / n	y / n	y / n
Lifestyle Modifications	y / n	y / n	y / n

### 3.3 Episodic instances of reflux (as opposed to the long-term disease GERD)

Although our basic model of GERD does not rely on minute by minute monitoring, we can trigger episodic symptoms, which can be important for diagnostic purposes. The physiological cause-effect (rather than clinically oriented) scripts cover the following:

- When a bolus (going down) or chyme (coming up) is located in a given esophageal segment, make the pH of the mucosa of that segment the same as the pH of the bolus or chyme for 3 minutes.
- If the pH of the mucosa is  $< 4$ , then a pH receptor is fired. In 15% of cases, when a pH receptor is fired, the patient will experience pain of greater than .2.
- If the mucosa is the theme of inflammation, then a pain-receptor is fired. In 15% of cases, when a pain receptor is fired, the patient will experience pain of greater than .2.
- We change the pH of chyme when patients take H2-blockers or PPIs so that pain during reflux does not occur.
- We permit reflux only if a massive tumor is not blocking T10 (otherwise no liquid could make it up the esophagus)

## Appendix 3-F

- We generate a drop in LES pressure if a person ingests some food/drink that is a GERD irritant for him. The drop is sufficient to make the LES pressure less than the gastric pressure (regardless of what the person's gastric pressure is), therefore causing reflux.
- We generate dysphagia based on lumen blockage and inflammation of the mucosa, as follows:

```
(if ((bolus (state-of-matter (value solid))) and (lumen-of-esophagus:R_destination (diameter (<> 1.4 1.8))))  
  then (dysphagia (domain (value human)) (range .2)))  
(if ((bolus (state-of-matter (value solid))) and (lumen-of-esophagus:R_destination (diameter (<> 1.0 1.39))))  
  then (dysphagia (domain (value human)) (range .5)))  
(if ((bolus (state-of-matter (value solid))) and (lumen-of-esophagus:R_destination (diameter (< .99))))  
  then (motion-event:R_bolus (epistemic 0))) ;; THIS event – can I just say ‘then (epistemic 0)’ with the motion-event implied?  
(if ((bolus (state-of-matter (value liquid))) and (lumen-of-esophagus:R_destination (diameter (<> .3 .99))))  
  then (dysphagia (domain (value human)) (range .5)))  
(if ((bolus (state-of-matter (value liquid))) and (lumen-of-esophagus:R_destination (diameter (< .3))))  
  then (motion-event:R_bolus (epistemic 0)))  
(if (mucosa-of-esophagus:R_destination (theme-of (or inflammation erosion ulcer)))  
  then (dysphagia (domain (value human)) (range .1))))
```

We have some physiological stubs as well, which we can expand later if needed:

- inflammation, erosion and ulcer all include: (have-event-as-part die-animate-part@mucosa-cells)
- the healing of inflammation, erosion and ulcer all include (have-event-as-part repopulate@mucosa-T10)
- we have a concept transient-lower-LES-relaxation, but we don't actually exploit it

### **Exploiting the system's knowledge of episodic reflux, in contrast to the general disease progression.**

We include among the Office Tests the request for the patient to drink/eat potential GERD-irritating substances (e.g., caffeine, chocolate) and report if this brought on symptoms. If so, following up by ingesting Tums should take away the symptoms. The problem with this model, however, is that most patients do not experience pain with every incidence of reflux.

## Appendix 3-F

### 3.4 GERD Test Results Repository

The tests carried out to establish GERD are pH monitoring, esophagogastroduodenoscopy and esophagogastroduodenoscopy with biopsy. The latter is done when a tumor is present. Test results are, as always, generated based on physiological prerequisites recorded in the TRR. (For a description of the TRR, see the document describing achalasia.)

(total-time-in-reflux-result

(definition “indicates how many hours out of 24 the person has a  $\text{ph} < 4$  in the esophagus”)

(comments “”)

(test-result-as-entity

(specialists-interpretation “Total time in reflux: ” (value ^\$var1) “ hours”)

(test-result-of ph-monitoring)

(test-result-content

(total-time-in-reflux (value ^\$var1))))))

(demeester-score-result

(definition “The DeMeester score is a composite score made up of all the factors quantified in the ph probe. These include percent time reflux upright, percent time reflux supine, percent time reflux total, episodes greater than 5 min, longest episode and total number of episodes. For each of these parameters there is a normal value, your value and a score given which depends on how much you vary from normal. These scores are added up and you get a composite demeester score. A normal DeMeester score is ( $< 14.7$ ), an abnormal DeMeester score is ( $> 14.7$ )”)

(comments “”)

(test-result-as-entity

(specialists-interpretation “DeMeester score :” ^\$var1)

(test-result-of ph-monitoring)

(test-result-content

(demeester-score (value ^\$var1))))))

## Appendix 3-F

```
(symptom-correlation-result
  (definition "indicates symptom correlation between GERD symptoms and ph-levels; if > 50%, that's normal")
  (comments "")
  (test-result-as-entity
    (specialists-interpretation "Symptom correlation: " ^$var1)
    (test-result-of ph-monitoring)
    (test-result-content
      (symptom-correlation (value ^$var1))))))
```

```
(esophageal-inflammation
  (definition "")
  (comments "")
  (test-result-as-entity
    (specialists-interpretation "There is inflammation in lower esophagus")
    (test-result-of esophagogastroduodenoscopy)
    (test-result-content
      (inflammation
        (theme mucosa-of-esophagus
          (part-of-object (value T10-segment-of-esophagus)))))))
```

```
(erosion-in-esophagus
  (definition "")
  (comments "")
  (test-result-as-entity
    (specialists-interpretation "There is erosion in the distal esophagus")
    (test-result-of esophagogastroduodenoscopy)
    (test-result-content
      (erosion
        (theme mucosa-of-esophagus
```

## Appendix 3-F

(part-of-object (value T10-segment-of-esophagus))))))

(ulcer-in-esophagus  
  (definition "")  
  (comments "")  
  (test-result-as-entity  
    (specialists-interpretation "There is an ulcer in the distal esophagus")  
    (test-result-of **esophagogastroduodenoscopy**)  
    (test-result-content  
      (ulcer  
        (theme mucosa-of-esophagus  
          (part-of-object (value T10-segment-of-esophagus  
            (part-of-object (value esophagogastroduodenoscopy.experiencer)))))) )

(barretts-esophagus-test-result  
  (definition "")  
  (comments "")  
  (test-result-as-entity  
    (specialists-interpretation "Barrett's esophagus")  
    (test-result-of **esophagogastroduodenoscopy-with-biopsy**)  
    (test-result-content  
      (barretts-esophagus))))

(peptic-stricture-test-result  
  (definition "")  
  (comments "")  
  (test-result-as-entity



## Appendix 3-F

(specialists-interpretation “peptic stricture”  
(test-result-of **esophagogastroduodenoscopy**)  
(test-result-content  
(peptic-stricture))))

(irregular-narrowing-of-distal-esophagus  
(definition “”)  
(comments “”)  
(test-result-as-entity  
(specialists-interpretation “There is an irregular narrowing of the distal esophagus. Suspicion of esophageal cancer.” “There is a  
several cm irregular narrowing of the distal esophagus with shouldering; possible esophageal cancer”)  
(test-result-of **barium-swallow**)  
(test-result-content  
(tumor (location distal-esophagus))))))

(smooth-narrowing-of-distal-esophagus  
(definition “”)  
(comments “note for MJ: seen for peptic stricture; not seen for Barrett’s – with Barrett’s, there’s no change in luminal size”)  
(test-result-as-entity  
(specialists-interpretation “There is smooth narrowing of the distal esophagus”)  
(test-result-of **esophagogastroduodenoscopy**)  
(test-result-content  
(diameter  
(domain t10-segment-of-esophagus)  
(range (< 2))))))

(tumor-in-esophagus  
(definition “”))

## Appendix 3-F

(comments “”)  
(test-result-as-entity  
    (specialists-interpretation “There is a tumor in the distal esophagus”)  
    (test-result-of **esophagogastroduodenoscopy**)  
    (test-result-content  
        (tumor  
            (location distal-esophagus))))))

## Authoring a GERD Patient, Part I (for all GERD patients)

Fill in the orange cells to create your patient.

Your patient must have either a low basic LES pressure or transient relaxations in order to have GERD.

<i>Basic Data Table</i>	
Last name	
First name	
Middle name	
Age	
Gender	
Race	
Basic LES pressure (0 – 35)	
Transient LES relaxations? (yes/no)	
GERD-irritating lifestyle Habits? (yes/no)	
Basic gastric pressure (0 – 10 mmHg)	

<a href="#">gerd-sensitive-food table</a>	Does ingesting the following types of things
---	--

## Appendix 3-F

	bring on GERD symptoms in your VP? (yes/no)
chocolate	
caffeine	
mints	
alcohol	
fatty food	
large meal	

There are 3 “mega-scenarios” for GERD:

1. inflammation – erosion – ulcer – peptic stricture
2. inflammation – Barrett’s esophagus – tumor
3. proximal GERD [not currently active]

Each of these has subtypes depending on how far along the progression the patient is predisposed to go. We have delineated 7 actual scenarios. Each one is supplied with the implicit predispositions of the patients who experience that scenario. The scenarios indicate the progression and end point of the disease: e.g., inflammation means “just inflammation – no further disease possible in this patient no matter how long he/she has GERD”.

Choose 1 scenario for your patient, then fill in the applicable follow-up questions about only that scenario.

Basic GERD Scenario Table	Type ‘yes’ to select	Scenarios	Properties	Values
1		inflammation	<p>predisposition-to-erosion</p> <p>predisposition-to-ulcer</p> <p>predisposition-to-stricture</p> <p>predisposition-to-Barretts</p> <p>predisposition-to-tumor</p> <p>predisposition-to-proximal-reflux</p>	<p>no</p> <p>no</p> <p>no</p> <p>no</p> <p>no</p> <p>no</p>
2		inflammation	predisposition-to-erosion	yes

### Appendix 3-F

		erosion	predisposition-to-ulcer predisposition-to-stricture predisposition-to-Barretts predisposition-to-tumor predisposition-to-proximal-reflux	no no no no no
3		inflammation erosion ulcer	predisposition-to-erosion predisposition-to-ulcer predisposition-to-stricture predisposition-to-Barretts predisposition-to-tumor predisposition-to-proximal-reflux	yes yes no no no no
4		inflammation erosion ulcer peptic-stricture	predisposition-to-erosion predisposition-to-ulcer predisposition-to-stricture predisposition-to-Barretts predisposition-to-tumor predisposition-to-proximal-reflux	yes yes yes no no no
5		inflammation Barrett's	predisposition-to-erosion predisposition-to-ulcer predisposition-to-stricture predisposition-to-Barretts predisposition-to-tumor predisposition-to-proximal-reflux	no no no yes no no
6		inflammation Barrett's tumor	predisposition-to-erosion predisposition-to-ulcer predisposition-to-stricture predisposition-to-Barretts predisposition-to-tumor predisposition-to-proximal-reflux	no no no yes yes no
7	[not active]	proximal_reflux	predisposition-to-erosion	no

## Appendix 3-F

			predisposition-to-ulcer	no
			predisposition-to-stricture	no
			predisposition-to-Barretts	no
			predisposition-to-tumor	no
			predisposition-to-proximal-reflux	yes

Now open the file that corresponds to your scenario and answer the follow-up questions.

For reasons of space, we present the follow-up questions for only two of the scenarios: “to peptic stricture” and “to ulcer”, which subsume the other scenarios.

### Scenario 4 (“to peptic stricture”) Follow-up Questions

Please provide the patient’s name exactly as on the first form.

Last name	
First name	
Middle name	

## Appendix 3-F

Indicate the rate of progression of GERD from inflammation to erosion, erosion to ulcer, and ulcer to peptic stricture by indicating the duration of the inflammation stage (t1) and the erosion stage (t2), the ulcer stage (t3) and the stricture stage (t4).

t-times table “to stricture”	length in days
t1	
t2	
t3	
t4	

In each table below, the ranges of values in white cells act as a guide. In the orange cell below each white cell, fill in an actual value (not a range) for your patient at the *start* of the given time period. Interpolation of changes throughout a time period, based on the start value of the next time period, will be done by the simulator. If you do not fill in a value explicitly, the default value will be used (written in parentheses next to the range). In some cases, no default has been provided and you must select a value explicitly. Hit the Tab key to jump to next cell in given row.

You need to select values for total time in reflux and the corresponding DeMeester score. A crib is provided for general orientation but it is not binding: you may create other associations as well. Note: In general, patients with mild reflux (~ 20%) may be managed with adherence to lifestyle modifications. One would assume that those with reflux less than or equal to 1.5 hrs and a DeMeester score of less than or equal to 20 would fit this profile. This expectation is shown in the table by the contrast between “Total time in reflux with bad habits” and “Total time in reflux with bad habits removed”. If your patient has no bad habits, then these should be the same.

DeMeester Crib

Time in Reflux (hours)	FYI: converted to percent time in reflux...	DeMeester Score
0	0	0
.96	4%	7
1.2	5%	10
1.44	6%	18
1.5	6.25%	20
1.68	7%	25
1.92	8%	32

## Appendix 3-F

2.4	10%	40
2.88	12%	48
3.0	12.5%	50
3.6	15%	60
4.5	18.75%	70
4.8	20%	80
6.0	25%	120

Basic physio table “to stricture”	t1 “inflammation”	t2 “erosion”	t3 “ulcer”	t4 “peptic-stricture”
Total time in reflux, incl. bad habits <b>in HOURS (use crib)</b>	1.2 – 1.44 (1.44) [5 - 6% [6]]	1.68 – 2.4 (1.92) [7– 10% [8]]	2.4 – 3.6 (2.88) [10 – 15% [12]]	1.68 – 3.6 (3.6) [7 – 15% [15]]
<i>TTR, bad habits, actual</i>				
DeMeester, incl. bad habits: <b>use crib</b>	10 – 18 (18)	25 – 40 (32)	40 – 60 (48)	25 – 60 (60)
<i>DeM., bad habits, actual</i>				
Total time in reflux when bad habits removed: <b>use crib</b>	.96 – 1.44 (.96) [4 – 6 % [4]]	1.68 – 2.4 (1.92) [7– 10% [8]]	2.4 – 3.6 (2.88) [10 – 15% [12]]	1.68 – 3.6 (3.6) [7 – 15% [15]]
<i>TTR, cleaned up, actual</i>				
DeMeester, when bad	7 – 18 (7)	25 – 40 (32)	40 – 60 (48)	25 – 60 (60)

### Appendix 3-F

habits removed: <b>use crib</b>				
<i>DeM., cleaned up, actual</i>				
Length of erosion (cm.)		S: .001 E: .5 – 4 (2)		
<i>Erosion length, actual</i>				
Diameter of erosion (mm.)		S: .001 E: .5 - 3 (1)		
<i>Erosion diameter, actual</i>				
Number of erosions		S: 1 - 4 E: 1 – 4 (2)		
<i># erosions, actual</i>				
Depth of ulcer (mm.)			S: 1 E: 1 – 3 (2)	
<i>Ulcer depth, actual</i>				
Diameter of ulcer (mm.)			S: 4 E: 4-10 (5)	
<i>Ulcer diameter, actual</i>				
Number of ulcers			S: 1 - 4 E: 1 – 4 (2)	
<i># ulcers, actual</i>				



### Appendix 3-F

Diameter of lumen of T10 (cm.)				S: guy's default E: .5
<i>T10 diam., actual</i>				

Symptoms Table "to stricture"	t1 "inflammation"	t2 "erosion"	t3 "ulcer"	t4 "peptic-stricture"
Heartburn (pain location chest) frequency (#/day)	3 – 5 (4)	6 – 8 (7)	9 – 10 (9)	6 – 10 (7)
<i>heartburn frequency, actual</i>				
Heartburn severity	.3 – .5 (.4)	.6 – .8 (.7)	.9 – 1.0 (.9)	.6 – 1.0 (.7)
<i>heartburn severity, actual</i>				
symptom correlation (for ph monitoring)	0 – 1	0 - 1	0 – 1	0 – 1
<i>symptom correl., actual</i>				
regurgitation freq. (# per day)	3 – 5 (4)	6 – 8 (7)	9 – 10 (9)	6 – 10 (7)
<i>regurg., actual</i>				

Treatment table "to stricture"	t1 "inflammation"	t2 "erosion"	t3 "ulcer"	t4 "peptic-stricture"
<b>PPI</b>	effective / ineffective	effective / ineffective	effective / ineffective	effective / ineffective
<i>PPI, actual</i>				

## Appendix 3-F

<b>H2 Blocker</b>	effective / ineffective	effective / ineffective	effective / ineffective	effective / ineffective
<i>H2 Blocker, actual</i>				
<b>Lifestyle Modifications</b>	compliance/non-compliance	ineffective	ineffective	ineffective
<i>Lifestyle mods, actual</i>				
<b>LES diameter after esophageal dilation TTS</b>				1 - 3
<i>LES diam. post-TTS, actual</i>				

## Scenario 6 (“to tumor”) Follow-up Questions

Please provide the patient’s name exactly as on the first form.

Last name	
First name	
Middle name	

## Appendix 3-F

In each table below, the ranges of values in white cells act as a guide. In the orange cell below each white cell, fill in an actual value (not a range) for your patient. If you do not fill in a value explicitly, the default value will be used (written in parentheses next to the range). In some cases, no default has been provided and you must select a value explicitly. Hit the Tab key to jump to next cell in given row.

You need to select values for total time in reflux and the corresponding DeMeester score. A crib is provided for general orientation but it is not binding: you may create other associations as well. Note: In general, patients with mild reflux (~ 20%) may be managed with adherence to lifestyle modifications. One would assume that those with reflux less than or equal to 1.5 hrs and a DeMeester score of less than or equal to 20 would fit this profile. This expectation is shown in the table by the contrast between “Total time in reflux with bad habits” and “Total time in reflux with bad habits removed”. If your patient has no bad habits, then these should be the same.

DeMeester Crib

Time in Reflux (hours)	FYI: converted to percent time in reflux...	DeMeester Score
0	0	0
.96	4%	7
1.2	5%	10
1.44	6%	18
1.5	6.25%	20
1.68	7%	25
1.92	8%	32
2.4	10%	40
2.88	12%	48
3.0	12.5%	50
3.6	15%	60
4.5	18.75%	70
4.8	20%	80
6.0	25%	120

Basic physio table “to tumor”	t1 “inflammation”	t2 “Barrett’s”	t3 “tumor”
-------------------------------	-------------------	----------------	------------

### Appendix 3-F

Total time in reflux, incl. bad habits: <b>use crib</b> (hrs. per 24)	1.2 – 1.44 (1.44) [5 - 6% [6]]	1.68 – 2.4 (1.92) [7– 10% [8]]	1.68 – 2.4 (1.92) [7– 10% [8]]
<i>TTR, bad habits, actual</i>			
DeMeester, incl. bad habits: <b>use crib</b>	10 – 18 (18)	25 – 40 (32)	25 – 40 (32)
<i>DeM., bad habits, actual</i>			
Total time in reflux, bad habits removed: <b>use crib</b> (hrs. per 24)	.96 – 1.44 (.96) [4 – 6 % [4]]	1.68 – 2.4 (1.92) [7– 10% [8]]	1.68 – 2.4 (1.92) [7– 10% [8]]
<i>TTR, no bad h., actual</i>			
DeMeester, when bad habits removed: <b>use crib</b>	7 – 18 (7)	25 – 40 (32)	25 – 40 (32)
<i>DeM., no bad h., actual</i>			
size of residual lumen	<i>normal</i>	<i>normal</i>	S: 3 cm. E: 1 cm. over a period of 1 year (fixed)
<i>Tumor size, actual</i>			<i>calculated</i>

## Appendix 3-F

Symptoms Table “to tumor”	t1 “inflammation”	t2 “Barrett’s”	t3 “Tumor”
Dysphagia	0 - .3 (.1)	0 - .3 (.1)	<i>calculated</i>
<i>dysphagia actual</i>			<i>calculated</i>
Heartburn (pain location chest) frequency (#/day)	3 – 5 (4)	0 - 8 (4)	0 - 8 (4)
<i>heartburn frequency, actual</i>			
Heartburn severity	.3 – .5 (.4)	.1 - .8 (.4)	.1 - .8 (.4)
<i>heartburn severity, actual</i>			
symptom correlation (for ph monitoring)	0 - 1	0 - 1	0 - 1
<i>symptom correlation, actual</i>			
regurgitation	3 – 5 (4)	6 – 8 (7)	6 – 8 (7)
<i>regurgitation, actual</i>			

Treatment “to tumor”	t1 “inflammation”	t2 “Barrett’s”	t3 “tumor”
<b>PPI QD</b>	effective / ineffective	effective / ineffective	effective/ineffective
<i>PPI QD, actual</i>			
<b>PPI BID</b>	effective/ineffective	effective/ineffective	effective/ineffective

### Appendix 3-F

<i>PPI BID, actual</i>			
<b>H2 Blocker</b>	effective/ineffective effective / ineffective	effective/ineffective effective / ineffective	effective/ineffective ineffective/ineffective
<i>H2 Blocker, actual</i>			
<b>Lifestyle Modifications</b>	compliance/non- complianceeffective / ineffective	ineffective	ineffective
<i>Lifestyle mods, actual</i>			

# An Actual GERD Patient – Implemented and Tested

Fill in the orange cells to create your patient.

Your patient must have either a low basic LES pressure or transient relaxations in order to have GERD.

<i>Basic Data Table</i>	
Last name	Jimenez
First name	Jose
Middle name	Juan
Age	38
Gender	m
Race	Hispanic
Basic LES pressure (0 – 35)	10
Transient LES relaxations? (yes/no)	Y
GERD-irritating lifestyle Habits? (yes/no)	Y
Basic gastric pressure (0 – 10 mmHg)	2
Description (teaching goals)	GERD to peptic stricture; alcohol sensitivity; H2 blockers ineffective throughout; PPI ineffective at peptic stricture stage as well

gerd-sensitive-food table	Does ingesting the following types of things bring on GERD symptoms in your VP? (yes/no)
chocolate	n
caffeine	n
mints	n
alcohol	y
fatty food	n
large meal	n

## Appendix 3-F

There are 3 “mega-scenarios” for GERD:

1. inflammation – erosion – ulcer – peptic stricture
2. inflammation – Barrett’s esophagus – tumor
3. proximal GERD [not currently active]

Each of these has subtypes depending on how far along the progression the patient is predisposed to go. We have delineated 7 actual scenarios. Each one is supplied with the implicit predispositions of the patients who experience that scenario. The scenarios indicate the progression and end point of the disease: e.g., inflammation means “just inflammation – no further disease possible in this patient no matter how long he/she has GERD”.

Choose 1 scenario for your patient, then fill in the applicable follow-up questions about only that scenario.

Basic GERD Scenario Table	Type ‘yes’ to select	Scenarios	Properties	Values
1	n	inflammation	predisposition-to-erosion predisposition-to-ulcer predisposition-to-stricture predisposition-to-Barretts predisposition-to-tumor predisposition-to-proximal-reflux	no no no no no no
2	n	inflammation erosion	predisposition-to-erosion predisposition-to-ulcer predisposition-to-stricture predisposition-to-Barretts predisposition-to-tumor predisposition-to-proximal-reflux	yes no no no no no
3	n	inflammation erosion ulcer	predisposition-to-erosion predisposition-to-ulcer predisposition-to-stricture predisposition-to-Barretts predisposition-to-tumor predisposition-to-proximal-reflux	yes yes no no no no
4	y	inflammation erosion ulcer	predisposition-to-erosion predisposition-to-ulcer predisposition-to-stricture	yes yes yes



## Appendix 3-F

		peptic-stricture	predisposition-to-Barretts predisposition-to-tumor predisposition-to-proximal-reflux	no no no
5	n	inflammation Barrett's	predisposition-to-erosion predisposition-to-ulcer predisposition-to-stricture predisposition-to-Barretts predisposition-to-tumor predisposition-to-proximal-reflux	no no no yes no no
6	n	inflammation Barrett's tumor	predisposition-to-erosion predisposition-to-ulcer predisposition-to-stricture predisposition-to-Barretts predisposition-to-tumor predisposition-to-proximal-reflux	no no no yes yes no
7	[not active]	proximal_reflux	predisposition-to-erosion predisposition-to-ulcer predisposition-to-stricture predisposition-to-Barretts predisposition-to-tumor predisposition-to-proximal-reflux	no no no no no yes

Now open the file that corresponds to your scenario and answer the follow-up questions.

### Scenario 4 (“to peptic stricture”) Follow-up Questions Gerd 4

Please provide the patient's name exactly as on the first form.

Last name	Jimenez
First name	Jose
Middle name	Juan

## Appendix 3-F

Indicate the rate of progression of GERD from inflammation to erosion, erosion to ulcer, and ulcer to peptic stricture by indicating the duration of the inflammation stage (t1) and the erosion stage (t2), the ulcer stage (t3) and the stricture stage (t4).

t-times table “to stricture”	length in days
t1	90
t2	150
t3	240
t4	360

In each table below, the ranges of values in white cells act as a guide. In the orange cell below each white cell, fill in an actual value (not a range) for your patient at the *start* of the given time period. Interpolation of changes throughout a time period, based on the start value of the next time period, will be done by the simulator. If you do not fill in a value explicitly, the default value will be used (written in parentheses next to the range). In some cases, no default has been provided and you must select a value explicitly. Hit the Tab key to jump to next cell in given row.

You need to select values for total time in reflux and the corresponding DeMeester score. A crib is provided for general orientation but it is not binding: you may create other associations as well. Note: In general, patients with mild reflux (~ 20%) may be managed with adherence to lifestyle modifications. One would assume that those with reflux less than or equal to 1.5 hrs and a DeMeester score of less than or equal to 20 would fit this profile. This expectation is shown in the table by the contrast between “Total time in reflux with bad habits” and “Total time in reflux with bad habits removed”. If your patient has no bad habits, then these should be the same.

DeMeester Crib

Time in Reflux (hours)	FYI: converted to percent time in reflux...	DeMeester Score
0	0	0
.96	4%	7
1.2	5%	10
1.44	6%	18
1.5	6.25%	20
1.68	7%	25
1.92	8%	32
2.4	10%	40
2.88	12%	48
3.0	12.5%	50
3.6	15%	60
4.5	18.75%	70

## Appendix 3-F

4.8	20%	80
6.0	25%	120

Basic physio table “to stricture”	t1 “inflammation”	t2 “erosion”	t3 “ulcer”	t4 “peptic-stricture”
Total time in reflux, incl. bad habits <b>in HOURS (use crib)</b>	1.2 – 1.44 (1.44) [5 - 6% [6]]	1.68 – 2.4 (1.92) [7– 10% [8]]	2.4 – 3.6 (2.88) [10 – 15% [12]]	1.68 – 3.6 (3.6) [7 – 15% [15]]
<i>TTR, bad habits, actual</i>	1.25	1.77	3.2	3.3
DeMeester, incl. bad habits: <b>use crib</b>	10 – 18 (18)	25 – 40 (32)	40 – 60 (48)	25 – 60 (60)
<i>DeM., bad habits, actual</i>	15	39	47	57
Total time in reflux when bad habits removed: <b>use crib</b>	.96 – 1.44 (.96) [4 – 6 % [4]]	1.68 – 2.4 (1.92) [7– 10% [8]]	2.4 – 3.6 (2.88) [10 – 15% [12]]	1.68 – 3.6 (3.6) [7 – 15% [15]]
<i>TTR, cleaned up, actual</i>	1.2	1.7	3.1	3.4
DeMeester, when bad habits removed: <b>use crib</b>	7 – 18 (7)	25 – 40 (32)	40 – 60 (48)	25 – 60 (60)
<i>DeM., cleaned up, actual</i>	14	37	46	58
Length of erosion (cm.)		S: .001 E: .5 – 4 (2)		
<i>Erosion length, actual</i>		3		

## Appendix 3-F

Diameter of erosion (mm.)		S: .001 E: .5 - 3 (1)		
<i>Erosion diameter, actual</i>		3		
Number of erosions		S: 1 - 4 E: 1 - 4 (2)		
<i># erosions, actual</i>		4		
Depth of ulcer (mm.)			S: 1 E: 1 - 3 (2)	
<i>Ulcer depth, actual</i>			3	
Diameter of ulcer (mm.)			S: 4 E: 4-10 (5)	
<i>Ulcer diameter, actual</i>			8	
Number of ulcers			S: 1 - 4 E: 1 - 4 (2)	
<i># ulcers, actual</i>			4	
Diameter of lumen of T10 (cm.)				S: 1.9 E: .5
<i>T10 diam., actual</i>				.5

Symptoms Table “to stricture”	t1 “inflammation”	t2 “erosion”	t3 “ulcer”	t4 “peptic-stricture”
Heartburn (pain location chest) frequency (#/day)	3 - 5 (4)	6 - 8 (7)	9 - 10 (9)	6 - 10 (7)
<i>heartburn frequency, actual</i>	3	6	9	6
Heartburn severity	.3 - .5 (.4)	.6 - .8 (.7)	.9 - 1.0 (.9)	.6 - 1.0 (.7)
<i>heartburn severity, actual</i>	.4	.6	.9	.6

## Appendix 3-F

symptom correlation (for ph monitoring)	0 – 1	0 - 1	0 – 1	0 – 1
<i>symptom correl., actual</i>	6	6	6	6
regurgitation freq. (# per day)	3 – 5 (4)	6 – 8 (7)	9 – 10 (9)	6 – 10 (7)
<i>regurg., actual</i>	3	6	9	6

Treatment table “to stricture”	t1 “inflammation”	t2 “erosion”	t3 “ulcer”	t4 “peptic-stricture”
<b>PPI</b>	effective / ineffective	effective / ineffective	effective / ineffective	effective / ineffective
<i>PPI, actual</i>	<b>e</b>	<b>e</b>	<b>e</b>	<b>i</b>
<b>H2 Blocker</b>	effective / ineffective	effective / ineffective	effective / ineffective	effective / ineffective
<i>H2 Blocker, actual</i>	<b>i</b>	<b>i</b>	<b>i</b>	<b>i</b>
<b>Lifestyle Modifications</b>	ineffective	ineffective	ineffective	ineffective
<i>Lifestyle mods, actual</i>	<b>i</b>	<b>i</b>	<b>i</b>	<b>i</b>
<b>LES diameter after esophageal dilation TTS</b>				1 - 3
<i>LES diam. post-TTS, actual</i>				<b>2</b>